(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) **World Intellectual Property Organization**
International Bureau

(43) **International Publication Date**
3 April 2003 (03.04.2003)

PCT

(10) **International Publication Number**
WO 03/027143 A2

(51) **International Patent Classification**[7]: **C07K 14/47**

(21) **International Application Number:** PCT/JP02/09876

(22) **International Filing Date:** 25 September 2002 (25.09.2002)

(25) **Filing Language:** English

(26) **Publication Language:** English

(30) **Priority Data:**
60/324,261 25 September 2001 (25.09.2001) US
60/391,666 26 June 2002 (26.06.2002) US
2,399,569 23 August 2002 (23.08.2002) CA

(71) **Applicants** *(for all designated States except US)*: **JAPAN as represented by THE PRESIDENT OF THE UNIVERSITY OF TOKYO** [JP/JP]; 3-1, Hongo 7-chome, Bunkyo-ku, Tokyo 113-8654 (JP). **ONCOTHERAPY SCIENCE, INC.** [JP/JP]; 3-16-13, Shirokanedai, Minato-ku, Tokyo 108-0071 (JP).

(72) **Inventors; and**
(75) **Inventors/Applicants** *(for US only)*: **NAKAMURA, Yusuke** [JP/JP]; 17-33, Azamino 1-chome, Aoba-ku, Yokohama-shi, Kanagawa 225-0011 (JP). **FURUKAWA, Yoichi** [JP/JP]; 10-40, Miyazaki 3-chome, Miyamae-ku, Kawasaki-shi, Kanagawa 216-0033 (JP).

(74) **Agents: SHIMIZU, Hatsushi** et al.; Kantetsu Tsukuba Bldg. 6F, 1-1-1, Oroshi-machi, Tsuchiura-shi, Ibaraki 300-0847 (JP).

(81) **Designated States** *(national)*: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, OM, PH, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, YU, ZA, ZM, ZW.

(84) **Designated States** *(regional)*: ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, SK, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:
— *without international search report and to be republished upon receipt of that report*

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

WO 03/027143 A2

(54) **Title:** GENE AND PROTEIN RELATING TO HEPATOCELLULAR CARCINOMA

(57) **Abstract:** The present application provides a novel human gene *ZNFN3A1* whose expression is markedly elevated in a great majority of HCCs compared to corresponding non-cancerous liver tissues. The gene encodes a protein having a zinc finger domain as well as a SET domain and has been found to form a regulatory complex with RNA helicase and RNA polymerase.

DESCRIPTION

Gene and Protein Relating To Hepatocellular Carcinoma

Technical Field

The present invention relates to the field of biological science, more specifically to the field of cancer research. In particular, the present invention relates a novel protein, ZNFN3A1, involved in the proliferation mechanism of hepatocellular carcinoma cells. The proteins of the present invention can be used, for example, as target molecules for developing drugs against liver cancer.

Background Art

Hepatocellular carcinoma (HCC) is one of the most common cancers worldwide and its incidence is gradually increasing in Japan as well as in United States (Akriviadis EA, et al., Br J Surg. 1998 Oct;85(10):1319-31). Although recent medical advances have made great progress in diagnosis, a large number of patients with HCCs are still diagnosed at advanced stages and their complete cures from the disease remain difficult. In addition, since patients with hepatic cirrhosis or chronic hepatitis have a high risk to HCCs, they may develop multiple liver tumors, or new tumors even after complete removal of initial tumors. Therefore development of highly effective chemotherapeutic drugs and preventive strategies are matters of pressing concern.

Recent advances in molecular biology suggest that multi-step processes underlie hepatocarcinogenesis as they do the genesis and progression of colon tumors. These processes involve qualitative and quantitative alterations of various gene products. Genetic alterations were observed in tumor suppressor genes including TP53 and AXIN1, and in oncogenes such as c-myc, cyclinD1, and s-catenin, in a subset number of HCCs (Tanaka S, et al., Cancer Res. 1993 Jun 15;53(12):2884-7; Satoh S, et al., Nat Genet. 2000 Mar;24(3):245-50). Frequent loss of heterozygosity (LOH) has been reported in chromosomal regions 4q, 6q, 8p, 8q, 9p, 9q, 13q, 16p, 16q, and 17p, in hepatic tumor cells (Feitelson MA, et al., Oncogene 2002 Apr 11;21(16):2593-604). In addition to these genetic changes, altered expression of a number of genes such as TGF α play a role in hepatocarcingenesis (Lee GH, et al., Cancer Res 1992 Oct 1;52(19):5162-70).

Summary of the Invention

The present inventors previously analyzed genome-wide expression of genes in hepatocellular carcinomas and identified a number of genes that

appeared to be involved in hepatocarcinogenesis. We also disclosed genes associated with type of hepatitis virus in the patients, histological grade of the tumors, and status of their vessel invasion (Okabe H, et al., Cancer Res. 2001 Mar 1;61(5):2129-37). Although inactivation of some tumor suppressor genes, such as p53 and p16INK4A, has been identified, until now no known oncogene has been shown to be commonly activated in hepatocellular carcinoma.

The present application provides an isolated gene, *ZNFN3A1,* that is overexpressed in hepatocellular carcinomas and an oncoprotein has been identified having of zinc finger domain and SET domain. The expression of the novel gene, *ZNFN3A1*, is frequently up-regulated in HCCs, and appears to confer an oncogenic activity to cancer cells through a transactivating RNA polymerase II complex, which, in turn, enhances transcription of target genes, including *EGFR*. Thus, *ZNFN3A1* serves as a novel molecular target for HCCs.

An object of the present invention is to provide a novel protein involved in the proliferation mechanism of hepatocellular carcinoma cells and the gene encoding the protein, as well as methods for producing and using the same in the diagnosis and treatment of hepatocellular carcinoma (HCC).

The inventors set out to discover novel targets of drug discovery for the treatment of HCC. Accordingly, they analyzed expression profiles of HCCs using a genome-wide cDNA microarray. Reported herein is the identification of a novel human gene *ZNFN3A1* whose expression is markedly elevated in a great majority of HCCs compared to corresponding non-cancerous liver tissues. The *ZNFN3A1* cDNA consists of 1622 nucleotides that contain an open reading frame of 1284 nucleotides encoding a putative 428-amino acid protein with a zinc finger motif. Interestingly subcellular localization of ZNFN3A1 protein is altered during cell cycle progression or due to the density of cultured cells; it accumulates in the nucleus when cells are in middle to late S phase or cultured in sparse condition, while it localizes in the cytoplasm as well as nucleus when they are in other phases or grown in dense condition. Furthermore, ZNFN3A1 directly associates with a RNA helicase KIAA0054, and forms a complex with RNA polymerase II, which activates transcription of downstream genes including epidermal growth factor receptor (EGFR) through a direct binding of the complex with an element of "(C)CCCTCC(T)" in the 5' flanking region. Consistently, exogenous expression of *ZNFN3A1* into NIH3T3 cells conferred increased cell growth, while suppression of its expression with antisense S-oligonucleotides resulted in a significant growth-inhibition of hepatoma cells. These findings suggest that ZNFN3A1 renders

oncogenic activities to cancer cells by transcriptional activation of target genes including *EGFR* through a complex with RNA helicase and RNA polymerase II, and that inhibition of the activity of the complex could be a promising strategy for the treatment of HCC.

Accordingly, the present invention relates to novel ZNFN3A1 proteins involved in cell proliferation and the genes encoding them, as well as the production and the use the same. More specifically, the present invention provides the following:

The present application provides a novel human protein, ZNFN3A1, or a functional equivalent thereof, that promotes cell proliferation and the transcriptional activation of target genes. In a preferred embodiment, the ZNFN3A1 protein includes a putative 428-amino acid protein with a zinc finger motif encoded by the open reading frame of SEQ. ID. NO.1. The zinc finger domain (MYND) is positioned at codons 49-87 and the SET (Su 3-9, Enhancer-of-zeste, Trihorrax) domain is positioned at codons 117-246. The ZNFN3A1 protein preferably includes the amino acid sequence set forth in SEQ. ID. NO.2. The present application also provides an isolated protein encoded from at least a portion of the *ZNFN3A1* polynucleotide sequence, or polynucloetide sequences at least 15%, and more preferably at least 25% complementary to the sequence set forth in SEQ. ID. NO. 1.

The present invention further provides a novel human gene, *ZNFN3A1*, whose expression is markedly elevated in a great majority of HCCs as compared to corresponding non-cancerous liver tissues. The isolated *ZNFN3A1* gene includes a polynucleotide sequence as described in SEQ. ID. NO. 1. In particular, the *ZNFN3A1* cDNA includes 1622 nucleotides that contain an open reading frame of 1284 nucleotides. The present invention further encompasses polynucleotides which hybridize to and which are at least 15%, and more preferably at least 25% complementary to the polynucleotide sequence set forth in SEQ. ID. NO. 1, to the extent that they encode a ZNFN3A1 protein or a functional equivalent thereof. Examples of such polynucleotides are degenerates of SEQ. ID. NO. 1.

As used herein, an “isolated *ZNFN3A1* gene” is a polynucleotide the structure of which is not identical to that of any naturally occurring polynucleotide or to that of any fragment of a naturally occurring genomic polynucleotide spanning more than three separate genes. The term therefore includes, for example, (a) a DNA which has the sequence of part of a naturally occurring genomic DNA molecule in the genome of the organism in which it naturally occurs;

(b) a polynucleotide incorporated into a vector or into the genomic DNA of a prokaryote or eukaryote in a manner such that the resulting molecule is not identical to any naturally occurring vector or genomic DNA; (c) a separate molecule such as a cDNA, a genomic fragment, a fragment produced by polymerase chain reaction (PCR), or a restriction fragment; and (d) a recombinant nucleotide sequence that is part of a hybrid gene, i.e., a gene encoding a fusion polypeptide. Specifically excluded from this definition are polynucleotides of DNA molecules present in mixtures of different (i) DNA molecules, (ii) transfected cells, or (iii) cell clones; e.g., as these occur in a DNA library such as a cDNA or genomic DNA library.

Accordingly, in one aspect, the invention provides an isolated polynucleotide that encodes a polypeptide described herein or a fragment thereof. Preferably, the isolated polypeptide includes a nucleotide sequence that is at least 60% identical to the nucleotide sequence shown in SEQ ID NO: 1. More preferably, the isolated nucleic acid molecule is at least 65%, 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or more, identical to the nucleotide sequence shown in SEQ ID NO: 1. In the case of an isolated polynucleotide which is longer than or equivalent in length to the reference sequence, e.g., SEQ ID NO: 1, the comparison is made with the full length of the reference sequence. Where the isolated polynucleotide is shorter than the reference sequence, e.g., shorter than SEQ ID NO: 1, the comparison is made to segment of the reference sequence of the same length (excluding any loop required by the homology calculation).

Also provided is a transcription activation complex that includes a ZNFN3A1 protein and at least one co-activator thereof. In one embodiment, the co-activator may be an RNA helicase and/or an RNA polymerase. The RNA helicase may be RNA helicase KIAA0054. The RNA polymerase may be RNA polymerase II. In one form, the complex includes the ZNFN3A1 protein, RNA helicase and RNA polymerase II. The transcription activation complex may activate transcription of genes including epidermal growth factor receptor (EGFR) through a direct binding of the complex with an element of "(C)CCCTCC(T)" in the 5' flanking region of the *EGFR* gene.

The present application also provides a therapeutic agent for treating a cancer. The therapeutic agent can be described as the polynucleotide sequence set forth in (SEQ. ID. NO. 3). The therapeutic agent may also be described as at least a portion of the antisense S-oligonucleotides of the ZNFN3A1 polynucleotide

sequence shown and described in SEQ. ID. NO. 1. A suitable antisense S-oligonucleotide is 5'-GCGGGAGGATGGAGCC-3' (SEQ. ID. NO. 3). The therapeutic agent may be suitably used to treat hepatoma cells. The course of action of the therapeutic agent is desirably to inhibit growth of hepatoma cells. The therapeutic agent may be applied to mammals including humans and domesticated mammals.

An antibody that recognizes the ZNFN3A1 protein is also provided by the present application. In part, an antisense DNA, ribozyme, and RNAi(RNA interference) of the ZNFN3A1 gene is also provided.

Further, a method of screening for a candidate compound for an anti-cancer agent is provided. The method includes contacting the ZNFN3A1 polypeptide with candidate compounds, and selecting compounds that bind to the ZNFN3A1 polypeptide.

The ZNFN3A1 polypeptide may also be contacted with the candidate compounds in the presence of a co-activator under the suitable condition for the formation of the complex of ZNFN3A1 polypeptide and the co-activator thereof. The compounds that inhibit the formation of the complex may then be selected. The co-activator may include RNA helicase and/or RNA polymerase II.

The present invention further provides a method of screening for a candidate compound for an anti-cancer agent, wherein the method includes contacting the ZNFN3A1 polypeptide, a co-activator thereof, and a DNA containing the target sequence of the polypeptide with candidate compounds under the suitable condition for the formation of the complex of ZNFN3A1 polypeptide and the DNA, and selecting compounds that inhibit the formation of the complex. The target sequence is desirably a CBS sequence flanking the 5' region of *EGFR*.

Also provided is a method of screening for a candidate compound for an anti-cancer agent, wherein the method includes contacting the ZNFN3A1 polypeptide, a co-activator thereof, and a reporter gene with a transcriptional regulatory region recognized by the complex of the polypeptide and the co-activator with candidate compounds under the suitable condition for the expression of the reporter gene, and selecting compounds that inhibit the expression of the reporter gene.

The present invention further provides a method for diagnosis of cancer that includes determining an expression level of the ZNFN3A1 gene in biological sample of specimen, comparing the expression level of ZNFN3A1 gene with that in normal sample, and defining a high expression level of the ZNFN3A1 gene in

the sample as having a cancer. The cancer is suitably a hepatocellular carcinoma.

The present invention also provides a method of producing a protein by transfecting or transforming a host cell with a polynucloetide sequence encoding the ZNFN3A1 protein, and expressing the polynucloetide sequence.

It is to be understood that both the foregoing summary of the invention and the following detailed description are of a preferred embodiment, and not restrictive of the invention or other alternate embodiments of the invention.

Brief Description of the Drawings

Figures 1A - 1C depict the expression of A6681 and ZNFN3A1 in HCCs. Figure 1A depicts the relative expression ratios of A6681 in 20 primary HCCs examined by cDNA microarray. Its expression was significantly up-regulated in eleven of the twelve (91.7%) clinical HCCs that passed through the cutoff filter (both Cy3 and Cy5 signals greater than 25,000). Figure 1B presents photographs depicting the expression of HCC candidate genes shown by ethidium bromide staining of RT-PCR products (T, tumor tissue; N, normal tissue). Expression of GAPDH served as an internal control. Figure 1C is a photograph depicting multiple-tissue northern blot analysis of A6681 in various adult human tissues.

Figure 2A depicts the predicted protein structure and protein motifs of ZNFN3A1. The protein motifs of ZNFN3A1 were predicted by Simple Modular Architecture Research Tool (SMART). Figure 2B depicts the homology between the deduced amino acid sequence of ZNFN3A1 and AK010447. The zf-MYND [zinc finger protein (MYND domain containing)] domain (A), and SET [(Su (var) 3-9, Enhancer-of-zeste, Trithorax)] domain (B) showed 94% and 95% identity, respectively.

Figures 3A - 3R presents photographs depicting the subcellular localization of ZNFN3A1 observed by immunocytochemistry, particularly through fluorescent immunohistochemical staining. Figures 3A - 3C depict SNU475 cells transfected with plasmid DNA designed to express EGFP-tagged ZNFN3A1 (pEGFP-ZNFN3A1). Fluorescent micrographs of FITC (g, j, m, p), DAPI (h, k, n, q), merge (I, l, o, r). Figures 3D - 3F depict SNU475 cells transfected with plasmid DNA designed to express FLAG-tagged ZNFN3A1 (pFLAG-ZNFN3A1). Fluorescent micrographs of anti-FLAG (d), DAPI (e), merge (f). Figures 3G - 3R depict endogenous expression of ZNFN3A1 in SNU475 cells (g-l) and SNU423 cells (m-r). Fluorescent micrographs of anti-ZNFN3A1 (g, j, m, p), DAPI (h, k, n, q), merge (i, l, o, r). Cells were inoculated in the low concentration (1.25 x 10^4 cells/well) (g-i, m-

o) and high concentration (1.0 x 10^5 cells/well) (j-l, p-r).

Figures 4A and 4B depict the effect of cell cycle progression on the sub-cellular localization of ZNFN3A1. Figure 4A depicts the FACS analysis of Huh7 cells synchronized by aphidicolin. The cells were growth-arrested in G_1 phase by incubation with 7.5 µg/ml aphidicolin for 36 h and released from G_1 by removal of aphidicolin. FACS was performed 0, 4, 8, and 12 h later. Figure 4B presents photographs depicting the fluorescent immunocytochemical staining of endogenous ZNFN3A1 protein in Huh7 synchronized by aphidicolin.

Figures 5A - 5D depict the growth promoting effect of *ZNFN3A1* in NIH3T3. Figure 5A is a photograph depicting the results of a colony formation assay of ZNFN3A1 in NIH3T3 cells, comparing the expression of ZNFN3A1 in mock, NIH3T3-antisense ZNFN3A1, and NIH3T3-sense ZNFN3A1. Figure 5B depicts the number of colonies counted by electric densitometry. Colonies were counted by electric densitometry. Colony numbers are presented as mean ± SD of triplicate plates. A (*) denotes a significant difference ($p< 0.05$) as determined by a Fisher's protected least significant test. Figure 5C presents photographs depicting the growth induction of NIH3T3 cells by expressing *ZNFN3A1* stably. Expression of *ZNFN3A1* mRNA in stable-transfectant (NIH3T3-ZNFN3A1) cells determined by RT-PCR. Figure 5D depicts the time course of cell numbers measured by trypan blue staining method.

Figures 6A - 6E depicts the growth suppressive effect of antisense S-oligonucleotides designated to suppress *ZNFN3A1*. Figure 6A depicts the construct of sense (Se) or antisense (As) oligonucleotides designed to suppress *ZNFN3A1*. Figure 6B presents photographs depicting the expression of *ZNFN3A1* in SNU475 cells treated with either sense (Se) or antisense (As) oligonucleotides for 24 h and examined by RT-PCR and western blotting using anti-ZNFN3A1 antibody. Figure 6C presents photographs depicting the results of a colony formation assay using sense or antisense oligonucleotides against *ZNFN3A1* mRNA in Huh7, Alexander, SNU423, and SNU475 cells. Antisense oligonucleotides (As) suppressed growth. Figure 6D depicts the results of an MIT assay, which assesses cell viability, 72 hours after sense or antisense oligonucleotide treatment in Huh7, Alexander, SNU423, and SNU475 cells. Figure 6E depicts the results of FACS analysis, which assesses cell cycle, of Huh7 cells 72 hours after treatment with sense (Se) or antisense (As) oligonucleotides.

Figures 7A - 7C depict the interaction between ZNFN3A1 and RNA Helicase KIAA0054. In Figure 7A, the conserved domain of KIAA0054 and the

regions for binding to ZNFN3A1 are mapped. Figure 7B presents photographs depicting the results of a yeast two-hybrid experiment, wherein pAS2-1 containing *ZNFN3A1* were co-transfected into yeast strain AH109 with library vectors containing two different length of KIAA0054. The results show the confirmation of binding between ZNFN3A1 and KIAA0054 in yeast strain AH109. Figure 7C presents photographs depicting the confirmation of binding between ZNFN3A1 and KIAA0054 in mammalian cells. Lysates from HeLa cells were immunoprecipitated with anti-FLAG or anti-HA antibody. The immunoprecipitates were analysed by immunoblotting with anti-HA or anti-FLAG antibody. Lysate was directly analysed by immunoblotting as a control.

Figure 8 presents photographs demonstrating that the interaction of ZNFN3A1 with KIAA0054 is mediated by SET domain and C-terminal region of KIAA0054. Deletion constructs of ZNFN3A1 were analysed for their ability to interact with KIAA0054 C-terminal region in the two-hybrid system.

Figure 9 presents photographs demonstrating that RNA helicase KIAA0054 associates with ZNFN3A1 and RNA polymerase II *in vivo*. Cellular extracts were prepared from HeLa cells transfected with 8 μg of the pFLAG-CMV-ZNFN3A1 (Full) and the pCMV-HA-KIAA0054 (Full length) expression vector. The extracts was immunoprecipitated with anti-RNA polymerase II antibody, anti-HA antibody or anti-FLAG antibody. The immunoprecipitates were analysed by immunoblotting with anti-RNA polymerase II antibody, anti-HA antibody or anti-FLAG antibody. Lysate was directly analysed by immunoblotting as a control.

Figures 10A and 10B demonstrate that candidate down-stream genes regulated by ZNFN3A1. Figure 10A depicts the sequence of oligonucleotides isolated by binding and amplification reactions with GST fusion proteins containing full length *ZNFN3A1*. The sequences of the core containing the random nucleotide region are shown. Figure 10B presents photographs depicting the reverse transcription analysis of extended transcripts of several genes in COS7 stable transformant that expressed ZNFN3A1 exogenously.

Figure11 (A) Schematic presentation of various reporter plasmids containing putative ZNFN3A1-binding elements in the 5' flanking region of *EGFR*. Nucleotide positions relative to the putative transcription-initiating site are indicated by plus or minus numbers. (B) Assay of the *EGFR* promoter in SNU475 cells using the indicated reporter plasmids. *Bars*, SD. *, a significant difference ($p < 0.05$) as determined by a Fisher's protected least-significant test.

Figure 12 is an illustration of complex formed between ZNFN3A1, RNA

helicase, and RNA polymerase II to regulate transcription of gene.

Detailed Description of the Invention

The words "a", "an", and "the" as used herein mean "at least one" unless otherwise specifically indicated.

The present application identifies a novel human gene *ZNFN3A1* whose expression is markedly elevated in HCCs compared to corresponding non-cancerous liver tissues. The *ZNFN3A1* cDNA consists of 1622 nucleotides that contain an open reading frame of 1284 nucleotides as set forth in SEQ. ID. NO. 1. The open reading frame encodes a putative 428-amino acid protein with a zinc finger motif. This protein has been named ZNFN3A1 (zinc finger protein, subfamily 3A (MYND containing), 1) by a nomenclature committee. Furthermore, ZNFN3A1 directly associates with a RNA helicase KIAA0054, and forms a complex with RNA polymerase II, which activates transcription of downstream genes including epidermal growth factor receptor (EGFR) through a direct binding of the complex with an element of "(C)CCCTCC(T)" in the 5' flanking region of the *EGFR* gene. Consistently, exogenous expression of *ZNFN3A1* into NIH3T3 cells conferred increased cell growth, while suppression of its expression with antisense S-oligonucleotides resulted in a significant growth-inhibition of hepatoma cells. These findings suggest that ZNFN3A1 renders oncogenic activities to cancer cells by transcriptional activation of target genes including *EGFR* through a complex with RNA helicase and RNA polymerase II, and that inhibition of the activity of the complex could be a promising strategy for the treatment of HCC.

The present invention encompasses novel human gene *ZNFN3A1*, including a polynucleotide sequence as described in SEQ. ID. NO. 1, as well as degenerates and mutants thereof, to the extent that they encode a ZNFN3A1 protein, including the amino acid sequence set forth in SEQ. ID. NO.2 or its functional equivalent. Examples of proteins functionally equivalent to ZNFN3A1 include, for example, homologous proteins of other organisms corresponding to the human ZNFN3A1 protein, as well as mutants of human ZNFN3A1 proteins.

In the present invention, the term "functionally equivalent" means that the subject protein has the activity to promote cell proliferation like ZNFN3A1 proteins and to confer oncogenic activity to cancer cells by forming a transactivating complex with an RNA helicase or an RNA polymerase or both, which, in turn, enhances the transcription of target genes, such as EGFR . Whether the subject protein has a cell proliferation activity or not can be judged by introducing the DNA

encoding the subject protein into a cell, such as NIH3T3, expressing the protein, and detecting promotion of proliferation of the cells or increase in colony forming activity. The ability of a subject protein to form a complex with either an RNA polymerase or an RNA helicase or both may be assayed by co-immunoprecipitation, such as those described in the Examples below. The enhancement of EGFR transcription may be further assayed using reporter plasmids such as those described in the Examples below.

Methods for preparing proteins functionally equivalent to a given protein are well known by a person skilled in the art and include known methods of introducing mutations into the protein. For example, one skilled in the art can prepare proteins functionally equivalent to the human ZNFN3A1 protein by introducing an appropriate mutation in the amino acid sequence of the human ZNFN3A1 protein by site-directed mutagenesis (Hashimoto-Gotoh, T. et al. (1995), Gene 152, 271-275; Zoller, MJ, and Smith, M. (1983), Methods Enzymol. 100, 468-500; Kramer, W. et al. (1984), Nucleic Acids Res. 12, 9441-9456; Kramer W, and Fritz HJ. (1987) Methods. Enzymol. 154, 350-367; Kunkel, TA (1985), Proc. Natl. Acad. Sci. USA. 82, 488-492; Kunkel (1988), Methods Enzymol. 85, 2763-2766). Amino acid mutations can occur in nature, too. The protein of the present invention includes those proteins having the amino acid sequences of the human ZNFN3A1 protein in which one or more amino acids are mutated, provided the resulting mutated proteins are functionally equivalent to the human ZNFN3A1 protein. The number of amino acids to be mutated in such a mutant is generally 10 amino acids or less, preferably 6 amino acids or less, and more preferably 3 amino acids or less.

Mutated or modified proteins, proteins having amino acid sequences modified by deleting, adding and/or replacing one or more amino acid residues of a certain amino acid sequence, have been known to retain the original biological activity (Mark, D. F. et al., Proc. Natl. Acad. Sci. USA (1984) 81, 5662-5666, Zoller, M. J. & Smith, M., Nucleic Acids Research (1982) 10, 6487-6500, Wang, A. et al., Science 224, 1431-1433, Dalbadie-McFarland, G. et al., Proc. Natl. Acad. Sci. USA (1982) 79, 6409-6413).

The amino acid residue to be mutated is preferably mutated into a different amino acid in which the properties of the amino acid side-chain are conserved (a process known as conservative amino acid substitution). Examples of properties of amino acid side chains are hydrophobic amino acids (A, I, L, M, F, P, W, Y, V), hydrophilic amino acids (R, D, N, C, E, Q, G, H, K, S, T), and side chains having the following functional groups or characteristics in common: an aliphatic side-

chain (G, A, V, L, I, P); a hydroxyl group containing side-chain (S, T, Y); a sulfur atom containing side-chain (C, M); a carboxylic acid and amide containing side-chain (D, N, E, Q); a base containing side-chain (R, K, H); and an aromatic containing side-chain (H, F, Y, W). Note, the parenthetic letters indicate the one-letter codes of amino acids.

An example of a protein to which one or more amino acids residues are added to the amino acid sequence of human ZNFN3A1 protein (SEQ. ID. NO. 2) is a fusion protein containing the human ZNFN3A1 protein. Fusion proteins are, fusions of the human ZNFN3A1 protein and other peptides or proteins, and are included in the present invention. Fusion proteins can be made by techniques well known to a person skilled in the art, such as by linking the DNA encoding the human ZNFN3A1 protein of the invention with DNA encoding other peptides or proteins, so that the frames match, inserting the fusion DNA into an expression vector and expressing it in a host. There is no restriction as to the peptides or proteins fused to the protein of the present invention.

Known peptides that can be used as peptides that are fused to the protein of the present invention include, for example, FLAG (Hopp, T. P. et al., Biotechnology (1988) 6, 1204-1210), 6xHis containing six His (histidine) residues, 10xHis, Influenza agglutinin (HA), human c-myc fragment, VSP-GP fragment, p18HIV fragment, T7-tag, HSV-tag, E-tag, SV40T antigen fragment, lck tag, α-tubulin fragment, B-tag, Protein C fragment, and the like. Examples of proteins that may be fused to a protein of the invention include GST (glutathione-S-transferase), Influenza agglutinin (HA), immunoglobulin constant region, β-galactosidase, MBP (maltose-binding protein), and such.

Fusion proteins can be prepared by fusing commercially available DNA, encoding the fusion peptides or proteins discussed above, with the DNA encoding the protein of the present invention and expressing the fused DNA prepared.

An alternative method known in the art to isolate functionally equivalent proteins is, for example, the method using a hybridization technique (Sambrook, J. et al., Molecular Cloning 2nd ed. 9.47-9.58, Cold Spring Harbor Lab. Press, 1989). One skilled in the art can readily isolate a DNA having high homology with a whole or part of the DNA sequence (SEQ. ID. NO. 1) encoding the human ZNFN3A1 protein, and isolate functionally equivalent proteins to the human ZNFN3A1 protein from the isolated DNA. The proteins of the present invention include those that are encoded by DNA that hybridize with a whole or part of the DNA sequence encoding the human ZNFN3A1 protein and are functionally equivalent to the

human ZNFN3A1 protein. These proteins include mammal homologues corresponding to the protein derived from human or mouse (for example, a protein encoded by a monkey, rat, rabbit and bovine gene). In isolating a cDNA highly homologous to the DNA encoding the human ZNFN3A1 protein from animals, it is particularly preferable to use tissues from ovary or testis.

The condition of hybridization for isolating a DNA encoding a protein functionally equivalent to the human ZNFN3A1 protein can be routinely selected by a person skilled in the art. For example, hybridization may be performed by conducting prehybridization at 68°C for 30 min or longer using "Rapid-hyb buffer" (Amersham LIFE SCIENCE), adding a labeled probe, and warming at 68°C for 1 hour or longer. The following washing step can be conducted, for example, in a low stringent condition. A low stringent condition is, for example, 42°C, 2X SSC, 0.1% SDS, or preferably 50°C, 2X SSC, 0.1% SDS. More preferably, high stringent conditions is used. A high stringent condition is, for example, washing 3 times in 2X SSC, 0.01% SDS at room temperature for 20 min, then washing 3 times in 1x SSC, 0.1% SDS at 37°C for 20 min, and washing twice in 1x SSC, 0.1% SDS at 50°C for 20 min. However, several factors such as temperature and salt concentration can influence the stringency of hybridization and one skilled in the art can suitably select the factors to achieved the requisite stringency.

In place of hybridization, a gene amplification method, for example, the polymerase chain reaction (PCR) method, can be utilized to isolate a DNA encoding a protein functionally equivalent to the human ZNFN3A1 protein, using a primer synthesized based on the sequence information of the DNA (SEQ. ID. NO. 1) encoding the human ZNFN3A1 protein.

Proteins that are functionally equivalent to the human ZNFN3A1 protein encoded by the DNA isolated through the above hybridization techniques or gene amplification techniques, normally have a high homology to the amino acid sequence of the human ZNFN3A1 protein. "High homology" typically refers to a homology of 40% or higher, preferably 60% or higher, more preferably 80% or higher, even more preferably 95% or higher. The homology of a protein can be determined by following the algorithm in "Wilbur, W. J. and Lipman, D. J. Proc. Natl. Acad. Sci. USA (1983) 80, 726-730".

A protein of the present invention may have variations in amino acid sequence, molecular weight, isoelectric point, the presence or absence of sugar chains, or form, depending on the cell or host used to produce it or the purification method utilized. Nevertheless, so long as it has a function equivalent to that of the

human ZNFN3A1 protein (SEQ. ID. NO. 2) of the present invention, it is within the scope of the present invention.

The proteins of the present invention can be prepared as recombinant proteins or natural proteins, by methods well known to those skilled in the art. A recombinant protein can be prepared by inserting a DNA , which encodes the protein of the present invention (for example, the DNA comprising the nucleotide sequence of SEQ. ID. NO. 1), into an appropriate expression vector, introducing the vector into an appropriate host cell, obtaining the extract, and purifying the protein by subjecting the extract to chromatography, for example, ion exchange chromatography, reverse phase chromatography, gel filtration, or affinity chromatography utilizing a column to which antibodies against the protein of the present invention is fixed, or by combining more than one of aforementioned columns.

Also when the protein of the present invention is expressed within host cells (for example, animal cells and *E. coli*) as a fusion protein with glutathione-S-transferase protein or as a recombinant protein supplemented with multiple histidines, the expressed recombinant protein can be purified using a glutathione column or nickel column.

After purifying the fusion protein, it is also possible to exclude regions other than the objective protein by cutting with thrombin or factor-Xa as required.

A natural protein can be isolated by methods known to a person skilled in the art, for example, by contacting the affinity column, in which antibodies binding to the ZNFN3A1 protein described below are bound, with the extract of tissues or cells expressing the protein of the present invention. The antibodies can be polyclonal antibodies or a monoclonal antibodies.

The present invention also encompasses partial peptides of the protein of the present invention. The partial peptide has an amino acid sequence specific to the protein of the present invention and consists of at least 7 amino acids, preferably 8 amino acids or more, and more preferably 9 amino acids or more. The partial peptide can be used, for example, for preparing antibodies against the protein of the present invention, screening for a compound that binds to the protein of the present invention, and screening for accelerators or inhibitors of the protein of the present invention.

A partial peptide of the invention can be produced by genetic engineering, by known methods of peptide synthesis, or by digesting the protein of the invention with an appropriate peptidase. For peptide synthesis, for example, solid phase

synthesis or liquid phase synthesis may be used.

The isolated ZNFN3A1 protein includes a putative 428-amino acid protein with a zinc finger motif encoded by the open reading frame of the *ZNFN3A1* polynucleotide sequence. The zinc finger domain (MYND) is positioned at codons 49-87 and the SET (Su 3-9, Enhancer-of-zeste, Trihorrax) domain is positioned at codons 117-246. As discussed in detail below, the ZNFN3A1 binding region resides in the SET domain. Therefore, the partial peptide of ZNFN3A1 preferably includes the SET domain.

Furthermore, the present invention provides DNA encoding the proteins of the present invention. The DNA of the present invention can be used for the *in vivo* or *in vitro* production of the protein of the present invention as described above, or can be applied to gene therapy for diseases attributed to genetic abnormality in the gene encoding the protein of the present invention. Any form of the DNA of the present invention can be used, so long as it encodes the protein of the present invention. Specifically, cDNA synthesized from the mRNA, genomic DNA, and chemically synthesized DNA can be used. The DNA of the present invention include a DNA comprising a given nucleotide sequences as well as its degenerate sequences, so long as the resulting DNA encodes a protein of the present invention.

The DNA of the present invention can be prepared by methods known to a person skilled in the art. For example, the DNA of the present invention can be prepared by: preparing a cDNA library from cells which express the protein of the present invention, and conducting hybridization using a partial sequence of the DNA of the present invention (for example, SEQ. ID. NO. 1) as a probe. A cDNA library can be prepared, for example, by the method described in Sambrook J. et al., Molecular Cloning, Cold Spring Harbor Laboratory Press (1989); alternatively, commercially available cDNA libraries may be used. A cDNA library can be also prepared by: extracting RNAs from cells expressing the protein of the present invention, synthesizing oligo DNAs based on the sequence of the DNA of the present invention (for example, SEQ. ID. NO. 1), conducting PCR by using the oligos as primers, and amplifying cDNAs encoding the protein of the present invention.

In addition, by sequencing the nucleotides of the obtained cDNA, the translation region encoded by the cDNA can be routinely determined, and the amino acid sequence of the protein of the present invention can be easily obtained. Moreover, by screening the genomic DNA library using the obtained cDNA as a

probe, the genomic DNA can be isolated.

More specifically, mRNAs may first be prepared from a cell, tissue, or organ (for example, ovary, testis, placenta, etc.) in which the protein of the invention is expressed. Known methods can be used to isolate mRNAs; for instance, total RNA may be prepared by guanidine ultracentrifugation (Chirgwin J. M. et al. Biochemistry 18:5294-5299 (1979)) or AGPC method (Chomczynski P. and Sacchi N. Anal. Biochem. 162:156-159 (1987)). In addition, mRNA may be purified from total RNA using mRNA Purification Kit (Pharmacia) and such or, alternatively, mRNA may be directly purified by QuickPrep mRNA Purification Kit (Pharmacia).

The obtained mRNA is used to synthesize cDNA using reverse transcriptase. cDNA may be synthesized by using a commercially available kit, such as the AMV Reverse Transcriptase First-strand cDNA Synthesis Kit (Seikagaku Kogyo). Alternatively, cDNA may be synthesized and amplified following the 5'-RACE method(Frohman M. A. et al. Proc. Natl. Acad. Sci. U.S.A. 85:8998-9002 (1988); Belyavsky A. et al. Nucleic Acids Res. 17:2919-2932 (1989)), which uses a primer and such, described herein, the 5'-Ampli FINDER RACE Kit (Clontech), and polymerase chain reaction (PCR).

A desired DNA fragment is prepared from the PCR products and ligated with a vector DNA. The recombinant vectors are used to transform *E. coli* and such, and a desired recombinant vector is prepared from a selected colony. The nucleotide sequence of the desired DNA can be verified by conventional methods, such as dideoxynucleotide chain termination.

The nucleotide sequence of a DNA of the invention may be designed to be expressed more efficiently by taking into account the frequency of codon usage in the host to be used for expression (Grantham R. et al. Nucleic Acids Res. 9:43-74 (1981)). The DNA of the present invention may be altered by a commercially available kit or a conventional method. For instance, the DNA may be altered by digestion with restriction enzymes, insertion of a synthetic oligonucleotide or an appropriate DNA fragment, addition of a linker, or insertion of the initiation codon (ATG) and/or the stop codon (TAA, TGA, or TAG).

Specifically, the DNA of the present invention encompasses the DNA comprising the nucleotide sequence encoding the zinc finger domain, positioned at codons 49-87 of SEQ. ID. NO.2 and the DET domain, positioned at codons 117-246 of SEQ. ID. NO.2

Furthermore, the present invention provides a DNA that hybridizes under

stringent conditions with a DNA having a nucleotide sequence of SEQ. ID. NO. 1, and encodes a protein functionally equivalent to the protein of the invention described above. One skilled in the art may appropriately choose stringent conditions. For example, low stringent condition can be used. More preferably, high stringent condition can be used. These conditions are the same as that described above. The hybridizing DNA above is preferably a cDNA or a chromosomal DNA.

The present invention also provides a vector into which a DNA of the present invention is inserted. A vector of the present invention is useful to keep a DNA of the present invention in host cell, or to express the protein of the present invention.

When *E. coli* is a host cell and the vector is amplified and produced in a large amount in *E. coli* (e.g., JM109, DH5α, HB101, or XL1Blue), the vector should have "ori" to be amplified in *E. coli* and a marker gene for selecting transformed *E. coli* (e.g., a drug-resistance gene selected by a drug such as ampicillin, tetracycline, kanamycin, chloramphenicol or the like). For example, M13-series vectors, pUC-series vectors, pBR322, pBluescript, pCR-Script, etc. can be used. In addition, pGEM-T, pDIRECT, and pT7 can also be used for subcloning and extracting cDNA as well as the vectors described above. When a vector is used to produce the protein of the present invention, an expression vector is especially useful. For example, an expression vector to be expressed in *E. coli* should have the above characteristics to be amplified in *E. coli*. When *E. coli*, such as JM109, DH5α, HB101, or XL1 Blue, are used as a host cell, the vector should have a promoter, for example, lacZ promoter (Ward et al., Nature (1989) 341, 544-546; FASEB J (1992) 6, 2422-2427), araB promoter (Better et al., Science (1988) 240, 1041-1043), or T7 promoter or the like, that can efficiently express the desired gene in *E. coli*. In that respect, pGEX-5X-1 (Pharmacia), “QIAexpress system” (Qiagen), pEGFP and pET (in this case, the host is preferably BL21 which expresses T7 RNA polymerase),for example, can be used instead of the above vectors.

Additionally, the vector may also contain a signal sequence for polypeptide secretion.. An exemplary signal sequence that directs the protein to be secreted to the periplasm of the *E. coli* is the pelB signal sequence (Lei, S. P. et al J. Bacteriol. (1987) 169, 4379). Means for introducing of the vectors into the target host cells include, for example, the calcium chloride method, and the electroporation method.

In addition to *E. coli*, for example, expression vectors derived from mammals (for example, pcDNA3 (Invitrogen) and pEGF-BOS (Nucleic Acids. Res.

1990, 18 (17), p5322), pEF, pCDM8), expression vectors derived from insect cells (for example, "Bac-to-BAC baculovirus expression system" (GIBCO BRL), pBacPAK8), expression vectors derived from plants (for example pMH1, pMH2), expression vectors derived from animal viruses (for example, pHSV, pMV, pAdexLcw), expression vectors derived from retroviruses (for example, pZIpneo), expression vector derived from yeast (for example, "Pichia Expression Kit" (Invitrogen), pNV11, SP-Q01), and expression vectors derived from *Bacillus subtilis* (for example, pPL608, pKTH50) can be used for producing the protein of the present invention.

In order to express the vector in animal cells, such as CHO, COS, or NIH3T3 cells, the vector should have a promoter necessary for expression in such cells, for example, the SV40 promoter (Mulligan et al., Nature (1979) 277, 108), the MMLV-LTR promoter, the EF1α promoter (Mizushima et al., Nucleic Acids Res. (1990) 18, 5322), the CMV promoter, and the like, and preferably a marker gene for selecting transformants (for example, a drug resistance gene selected by a drug (e.g., neomycin, G418)). Examples of known vectors with these characteristics include, for example, pMAM, pDR2, pBK-RSV, pBK-CMV, pOPRSV, and pOP13.

In addition, methods may be used to express a gene stably and, at the same time, to amplify the copy number of the gene in cells. For example, a vector comprising the complementary DHFR gene (for example pCHO I) may be introduced into CHO cells in which the nucleic acid synthesizing pathway is deleted, and then amplified by methotrexate (MTX). Furthermore, in case of transient expression of a gene, the method wherein a vector comprising a replication origin of SV40 (pcD, etc.) is transformed into COS cells comprising the SV40 T antigen expressing gene on the chromosome can be used.

A protein of the present invention obtained as above may be isolated from inside or outside (such as medium) of host cells, and purified as a substantially pure homogeneous protein. The term "substantially pure" as used herein in reference to a given polypeptide means that the polypeptide is substantially free from other biological macromolecules. The substantially pure polypeptide is at least 75% (e.g., at least 80, 85, 95, or 99%) pure by dry weight. Purity can be measured by any appropriate standard method, for example by column chromatography, polyacrylamide gel electrophoresis, or HPLC analysis. The method for protein isolation and purification is not limited to any specific method; in fact, any standard method may be used.

For instance, column chromatography, filter, ultrafiltration, salt precipitation, solvent precipitation, solvent extraction, distillation, immunoprecipitation, SDS-polyacrylamide gel electrophoresis, isoelectric point electrophoresis, dialysis, and recrystallization may be appropriately selected and combined to isolate and purify the protein.

Examples of chromatography include, for example, affinity chromatography, ion-exchange chromatography, hydrophobic chromatography, gel filtration, reverse phase chromatography, adsorption chromatography, and such (Strategies for Protein Purification and Characterization: A Laboratory Course Manual. Ed. Daniel R. Marshak et al., Cold Spring Harbor Laboratory Press (1996)). These chromatographies may be performed by liquid chromatography, such as HPLC and FPLC. Thus, the present invention provides for highly purified proteins prepared by the above methods.

A protein of the present invention may be optionally modified or partially deleted by treating it with an appropriate protein modification enzyme before or after purification. Useful protein modification enzymes include, but are not limited to, trypsin, chymotrypsin, lysylendopeptidase, protein kinase, glucosidase and so on.

The present invention provides an antibody that binds to the protein of the invention. The antibody of the invention can be used in any form, such as monoclonal or polyclonal antibodies, and includes antiserum obtained by immunizing an animal such as a rabbit with the protein of the invention, all classes of polyclonal and monoclonal antibodies, human antibodies, and humanized antibodies produced by genetic recombination.

A protein of the invention used as an antigen to obtain an antibody may be derived from any animal species, but preferably is derived from a mammal such as a human, mouse, or rat, more preferably from a human. A human-derived protein may be obtained from the nucleotide or amino acid sequences disclosed herein.

According to the present invention, the protein to be used as an immunization antigen may be a complete protein or a partial peptide of the protein. A partial peptide may comprise, for example, the amino (N)-terminal or carboxy (C)-terminal fragment of a protein of the present invention. Herein, an antibody is defined as a protein that reacts with either the full length or a fragment of a protein of the present invention.

A gene encoding a protein of the invention or its fragment may be inserted into a known expression vector, which is then used to transform a host cell as

described herein. The desired protein or its fragment may be recovered from the outside or inside of host cells by any standard method, and may subsequently be used as an antigen. Alternatively, whole cells expressing the protein or their lysates, or a chemically synthesized protein may be used as the antigen.

Any mammalian animal may be immunized with the antigen, but preferably the compatibility with parental cells used for cell fusion is taken into account. In general, animals of Rodentia, Lagomorpha, or Primates are used.

Animals of Rodentia include, for example, mouse, rat, and hamster. Animals of Lagomorpha include, for example, rabbit. Animals of Primates include, for example, a monkey of Catarrhini (old world monkey) such as *Macaca fascicularis*, rhesus monkey, sacred baboon, and chimpanzees.

Methods for immunizing animals with antigens are known in the art. Intraperitoneal injection or subcutaneous injection of antigens is a standard method for immunization of mammals. More specifically, antigens may be diluted and suspended in an appropriate amount of phosphate buffered saline (PBS), physiological saline, etc. If desired, the antigen suspension may be mixed with an appropriate amount of a standard adjuvant, such as Freund's complete adjuvant, made into emulsion, and then administered to mammalian animals. Preferably, it is followed by several administrations of antigen mixed with an appropriately amount of Freund's incomplete adjuvant every 4 to 21 days. An appropriate carrier may also be used for immunization. After immunization as above, serum is examined by a standard method for an increase in the amount of desired antibodies.

Polyclonal antibodies against the proteins of the present invention may be prepared by collecting blood from the immunized mammal examined for the increase of desired antibodies in the serum, and by separating serum from the blood by any conventional method. Polyclonal antibodies include serum containing the polyclonal antibodies, as well as the fraction containing the polyclonal antibodies may be isolated from the serum. Immunoglobulin G or M can be prepared from a fraction which recognizes only the protein of the present invention using, for example, an affinity column coupled with the protein of the present invention, and further purifying this fraction by using protein A or protein G column.

To prepare monoclonal antibodies, immune cells are collected from the mammal immunized with the antigen and checked for the increased level of desired antibodies in the serum as described above, and are subjected to cell

fusion. The immune cells used for cell fusion are preferably obtained from spleen. Other preferred parental cells to be fused with the above immunocyte include, for example, myeloma cells of mammalians, and more preferably myeloma cells having an acquired property for the selection of fused cells by drugs.

The above immunocyte and myeloma cells can be fused according to known methods, for example, the method of Milstein et al. (Galfre, G. and Milstein, C., Methods Enzymol. (1981) 73, 3-46).

Resulting hybridomas obtained by the cell fusion may be selected by cultivating them in a standard selection medium, such as HAT medium (hypoxanthine, aminopterin, and thymidine containing medium). The cell culture is typically continued in the HAT medium for several days to several weeks, the time being sufficient to allow all the other cells, with the exception of the desired hybridoma (non-fused cells), to die. Then, the standard limiting dilution is performed to screen and clone a hybridoma cell producing the desired antibody.

In addition to the above method, in which a non human animal is immunized with an antigen for preparing hybridoma, human lymphocytes such as those infected by EB virus may be immunized with a protein, protein expressing cells, or their lysates *in vitro*. Then, the immunized lymphocytes are fused with human-derived myeloma cells that are capable of indefinitely dividing, such as U266, to yield a hybridoma producing a desired human antibody that is able to bind to the protein can be obtained (Unexamined Published Japanese Patent Application No. (JP-A) Sho 63-17688).

The obtained hybridomas are subsequently transplanted into the abdominal cavity of a mouse and the ascites are extracted. The obtained monoclonal antibodies can be purified by, for example, ammonium sulfate precipitation, a protein A or protein G column, DEAE ion exchange chromatography, or an affinity column to which the protein of the present invention is coupled. The antibody of the present invention can be used not only for purification and detection of the protein of the present invention, but also as a candidate for agonists and antagonists of the protein of the present invention. In addition, this antibody can be applied to the antibody treatment for diseases related to the protein of the present invention. When the obtained antibody is to be administered to the human body (antibody treatment), a human antibody or a humanized antibody is preferable for reducing immunogenicity.

For example, transgenic animals having a repertory of human antibody genes may be immunized with an antigen selected from a protein, protein

expressing cells, or their lysates. Antibody producing cells are then collected from the animals and fused with myeloma cells to obtain hybridoma, from which human antibodies against the protein can be prepared (see WO92-03918, WO93-2227, WO94-02602, WO94-25585, WO96-33735, and WO96-34096).

Alternatively, an immune cell, such as an immunized lymphocyte, producing antibodies may be immortalized by an oncogene and used for preparing monoclonal antibodies.

Monoclonal antibodies thus obtained can be also recombinantly prepared using genetic engineering techniques (see, for example, Borrebaeck C. A. K. and Larrick J. W. Therapeutic Monoclonal Antibodies, published in the United Kingdom by MacMillan Publishers LTD (1990)). For example, a DNA encoding an antibody may be cloned from an immune cell, such as a hybridoma or an immunized lymphocyte producing the antibody, inserted into an appropriate vector, and introduced into host cells to prepare a recombinant antibody. The present invention also provides recombinant antibodies prepared as described above.

Furthermore, an antibody of the present invention may be a fragment of an antibody or modified antibody, so long as it binds to one or more of the proteins of the invention. For instance, the antibody fragment may be Fab, $F(ab')_2$, Fv, or single chain Fv (scFv), in which Fv fragments from H and L chains are ligated by an appropriate linker (Huston J. S. et al. Proc. Natl. Acad. Sci. U.S.A. 85:5879-5883 (1988)). More specifically, an antibody fragment may be generated by treating an antibody with an enzyme, such as papain or pepsin. Alternatively, a gene encoding the antibody fragment may be constructed, inserted into an expression vector, and expressed in an appropriate host cell (see, for example, Co M. S. et al. J. Immunol. 152:2968-2976 (1994); Better M. and Horwitz A. H. Methods Enzymol. 178:476-496 (1989); Pluckthun A. and Skerra A. Methods Enzymol. 178:497-515 (1989); Lamoyi E. Methods Enzymol. 121:652-663 (1986); Rousseaux J. et al. Methods Enzymol. 121:663-669 (1986); Bird R. E. and Walker B. W. Trends Biotechnol. 9:132-137 (1991)).

An antibody may be modified by conjugation with a variety of molecules, such as polyethylene glycol (PEG). The present invention provides for such modified antibodies. The modified antibody can be obtained by chemically modifying an antibody. These modification methods are conventional in the field.

Alternatively, an antibody of the present invention may be obtained as a chimeric antibody, between a variable region derived from nonhuman antibody and the constant region derived from human antibody, or as a humanized antibody,

comprising the complementarity determining region (CDR) derived from nonhuman antibody, the frame work region (FR) derived from human antibody, and the constant region. Such antibodies can be prepared by using known technology.

Antibodies obtained as above may be purified to homogeneity. For example, the separation and purification of the antibody can be performed according to separation and purification methods used for general proteins. For example, the antibody may be separated and isolated by the appropriately selected and combined use of column chromatographies, such as affinity chromatography, filter, ultrafiltration, salting-out, dialysis, SDS polyacrylamide gel electrophoresis, isoelectric focusing, and others (Antibodies: A Laboratory Manual. Ed Harlow and David Lane, Cold Spring Harbor Laboratory, 1988), but are not limited thereto.

A protein A column and protein G column can be used as the affinity column. Exemplary protein A columns to be used include, for example, Hyper D, POROS, and Sepharose F.F. (Pharmacia).

Exemplary chromatography, with the exception of affinity includes, for example, ion-exchange chromatography, hydrophobic chromatography, gel filtration, reverse-phase chromatography, adsorption chromatography, and the like (Strategies for Protein Purification and Characterization: A Laboratory Course Manual. Ed Daniel R. Marshak et al., Cold Spring Harbor Laboratory Press, 1996). The chromatographic procedures can be carried out by liquid-phase chromatography, such as HPLC, FPLC.

For example, measurement of absorbance, enzyme-linked immunosorbent assay (ELISA), enzyme immunoassay (EIA), radioimmunoassay (RIA), and/or immunofluorescence may be used to measure the antigen binding activity of the antibody of the invention. In ELISA, the antibody of the present invention is immobilized on a plate, protein of the invention is applied to the plate, and then a sample containing a desired antibody, such as culture supernatant of antibody producing cells or purified antibodies, is applied. Then, a secondary antibody that recognizes the primary antibody and is labeled with an enzyme, such as alkaline phosphatase, is applied, and the plate is incubated. Next, after washing, an enzyme substrate, such as *p*-nitrophenyl phosphate, is added to the plate, and the absorbance is measured to evaluate the antigen binding activity of the sample. A fragment of the protein, such as a C-terminal or N-terminal fragment, may be used as a protein. BIAcore (Pharmacia) may be used to evaluate the activity of the antibody according to the present invention.

The above methods allow for the detection or measurement of the protein of the invention, by exposing the antibody of the invention to a sample assumed to contain the protein of the invention, and detecting or measuring the immune complex formed by the antibody and the protein.

Because the method of detection or measurement of the protein according to the invention can specifically detect or measure a protein, the method may be useful in a variety of experiments in which the protein is used.

The present invention also provides a polynucleotide which hybridizes with the DNA encoding human ZNFN3A1 protein (SEQ. ID. NO.1) or the complementary strand thereof, and which comprises at least 15 nucleotides. The polynucleotide of the present invention is preferably a polynucleotide which specifically hybridizes with the DNA encoding the protein of the present invention. The term "specifically hybridize" as used herein, means that cross-hybridization does not occur significantly with DNA encoding other proteins, under the usual hybridizing conditions, preferably under stringent hybridizing conditions. Such polynucleotides include, probes, primers, nucleotides and nucleotide derivatives (for example, antisense oligonucleotides and ribozymes), which specifically hybridize with DNA encoding the protein of the invention or its complementary strand. Moreover, such polynucleotide can be utilized for the preparation of DNA chip.

The present invention includes an antisense oligonucleotide that hybridizes with any site within the nucleotide sequence of SEQ. ID. NO. 1. This antisense oligonucleotide is preferably against at least 15 continuous nucleotides of the nucleotide sequence of SEQ. ID. NO. 1. The above-mentioned antisense oligonucleotide, which contains an initiation codon in the above-mentioned at least 15 continuous nucleotides, is even more preferred.

Derivatives or modified products of antisense oligonucleotides can be used as antisense oligonucleotides. Examples of such modified products include lower alkyl phosphonate modifications such as methyl-phosphonate-type or ethyl-phosphonate-type, phosphorothioate modifications and phosphoroamidate modifications.

The term "antisense oligonucleotides" as used herein means, not only those in which the nucleotides corresponding to those constituting a specified region of a DNA or mRNA are entirely complementary, but also those having a mismatch of one or more nucleotides, as long as the DNA or mRNA and the antisense oligonucleotide can specifically hybridize with the nucleotide sequence of SEQ. ID.

NO. 1.

Such polynucleotides are contained as those having, in the "at least 15 continuous nucleotide sequence region", a homology of at least 70% or higher, preferably at 80% or higher, more preferably 90% or higher, even more preferably 95% or higher. The algorithm stated herein can be used to determine the homology. Such polynucleotides are useful as probes for the isolation or detection of DNA encoding the protein of the invention as stated in a later example or as a primer used for amplifications.

The antisense oligonucleotide derivatives of the present invention act upon cells producing the protein of the invention by binding to the DNA or mRNA encoding the protein, inhibiting its transcription or translation, promoting the degradation of the mRNA, and inhibiting the expression of the protein of the invention, thereby resulting in the inhibition of the protein's function.

An antisense oligonucleotide derivative of the present invention can be made into an external preparation, such as a liniment or a poultice, by mixing with a suitable base material which is inactive against the derivatives.

Also, as needed, the derivatives can be formulated into tablets, powders, granules, capsules, liposome capsules, injections, solutions, nose-drops and freeze-drying agents by adding excipients, isotonic agents, solubilizers, stabilizers, preservatives, pain-killers, and such. These can be prepared by following usual methods.

The antisense oligonucleotide derivative is given to the patient by directly applying onto the ailing site or by injecting into a blood vessel so that it will reach the site of ailment. An antisense-mounting medium can also be used to increase durability and membrane-permeability. Examples are, liposome, poly-L-lysine, lipid, cholesterol, lipofectin or derivatives of these.

The dosage of the antisense oligonucleotide derivative of the present invention can be adjusted suitably according to the patient's condition and used in desired amounts. For example, a dose range of 0.1 to 100 mg/kg, preferably 0.1 to 50 mg/kg can be administered.

The antisense oligonucleotide of the invention inhibits the expression of the protein of the invention and is thereby useful for suppressing the biological activity of the protein of the invention. Also, expression-inhibitors, comprising the antisense oligonucleotide of the invention, are useful in the point that they can inhibit the biological activity of the protein of the invention.

Moreover, the present invention provides a method of screening for a

compound that binds to the protein of the present invention by using the protein of the present invention. This screening method comprises the steps of: (a) contacting the protein of the present invention or a partial peptide thereof with a subject sample, (b) detecting the binding activity between the protein of the present invention or the partial peptide thereof and the subject sample, and (c) selecting a compound that binds to the protein of the present invention or the partial peptide thereof.

The protein of the present invention to be used for screening may be a recombinant protein or a protein derived from the nature, or a partial peptide thereof. Any subject sample, for example, cell extracts, cell culture supernatant, products of fermenting microorganism, extracts from marine organism, plant extracts, purified or crude proteins, peptides, non-peptide compounds, synthetic micromolecular compounds and natural compounds, can be used. The protein of the present invention to be contacted with a subject sample can be, for example, a purified protein, a soluble protein, a form bound to a carrier, or a fusion protein fused with other proteins.

As a method of screening for proteins, for example, that bind to the protein of the present invention using the protein of the present invention, many methods well known by a person skilled in the art can be used. Such a screening can be conducted by, for example, immunoprecipitation method, specifically, in the following manner. The gene encoding the protein of the present invention is expressed in animal cells and so on by inserting the gene to an expression vector for foreign genes, such as pSV2neo, pcDNA I, and pCD8. The promoter to be used for the expression may be any promoter that can be used commonly and include, for example, the SV40 early promoter (Rigby in Williamson (ed.), Genetic Engineering, vol. 3. Academic Press, London, p. 83-141 (1982)), the EF-1α promoter (Kim et al., Gene 91, p217-223 (1990)), the CAG promoter (Niwa et al. Gene 108, p. 193-200 (1991)), the RSV LTR promoter (Cullen Methods in Enzymology 152, p. 684-704 (1987)) the SRα promoter (Takebe et al., Mol. Cell. Biol. 8, p. 466 (1988), the CMV immediate early promoter (Seed and Aruffo Proc. Natl. Acad. Sci. USA 84, p. 3365-3369 (1987)), the SV40 late promoter (Gheysen and Fiers J. Mol. Appl. Genet. 1, p. 385-394 (1982)), the Adenovirus late promoter (Kaufman et al., Mol. Cell. Biol. 9, p. 946 (1989)), the HSV TK promoter and so on. The introduction of the gene into animal cells to express a foreign gene can be performed according to any methods, for example, the electroporation method (Chu G. et al. Nucl. Acids Res. 15, 1311-1326 (1987)), the calcium phosphate

method (Chen, C and Okayama, H. Mol. Cell. Biol. 7, 2745-2752 (1987)), the DEAE dextran method (Lopata, M. A. et al. Nucl. Acids Res. 12, 5707-5717 (1984)), Sussman, D. J. and Milman, G. Mol. Cell. Biol. 4, 1642-1643 (1985)), the Lipofectin method (Derijard, B. Cell 7, 1025-1037 (1994); Lamb, B. T. et al. Nature Genetics 5, 22-30 (1993): Rabindran, S. K. et al. Science 259, 230-234 (1993)), and so on. The protein of the present invention can be expressed as a fusion protein comprising a recognition site (epitope) of a monoclonal antibody by introducing the epitope of the monoclonal antibody, whose specificity has been revealed, to the N- or C- terminus of the protein of the present invention. A commercially available epitope-antibody system can be used (Experimental Medicine 13, 85-90 (1995)). Vectors which can express a fusion protein with, for example, β-galactosidase, maltose binding protein, glutathione S-transferase, green florescence protein (GFP) and so on by the use of its multiple cloning sites are commercially available.

A fusion protein prepared by introducing only small epitopes consisting of several to a dozen amino acids so as not to change the property of the protein of the present invention by the fusion is also reported. Epitopes, such as polyhistidine (His-tag), influenza aggregate HA, human c-myc, FLAG, Vesicular stomatitis virus glycoprotein (VSV-GP), T7 gene 10 protein (T7-tag), human simple herpes virus glycoprotein (HSV-tag), E-tag (an epitope on monoclonal phage), and such, and monoclonal antibodies recognizing them can be used as the epitope-antibody system for screening proteins binding to the protein of the present invention (Experimental Medicine 13, 85-90 (1995)).

In immunoprecipitation, an immune complex is formed by adding these antibodies to cell lysate prepared by using an appropriate detergent. The immune complex consists of the protein of the present invention, a protein comprising the binding ability with the protein, and an antibody. Immunoprecipitation can be also conducted by using antibodies against the protein of the present invention, besides using antibodies against the above epitopes. An antibody against the protein of the present invention can be prepared, for example, by introducing a gene encoding the protein of the present invention to an appropriate *E. coli* expression vector, expressing the gene in *E. coli*, purifying the expressed protein, and immunizing rabbits, mice, rats, goats, domestic fowls and such against the protein. The antibody can be also prepared by immunizing the above animals against a synthesized partial peptide of the protein of the present invention.

An immune complex can be precipitated, for example by Protein A

Sepharose or Protein G sepharose when the antibody is a mouse IgG antibody. If the protein of the present invention is prepared as a fusion protein with an epitope, such as GST, an immune complex can be formed in the same manner as in the use of the antibody against the protein of the present invention, by using a substance specifically binding to these epitopes, such as glutathione-Sepharose 4B.

Immunoprecipitation can be performed by following or according to, for example, the methods in the literature (Harlow, E. and Lane, D.: Antibodies pp. 511-552, Cold Spring Harbor Laboratory publications, New York (1988))

SDS-PAGE is commonly used for analysis of immunoprecipitated proteins and the bound protein can be analyzed by the molecular weight of the protein by using gels with an appropriate concentration. Since the protein bound to the protein of the present invention is difficult to detect by a common staining method, such as Coomassie staining or silver staining, the detection sensitivity for the protein can be improved by culturing cells in culture medium containing radioactive isotope, ^{35}S-methionine or ^{35}S-cystein, labeling proteins in the cells, and detecting the proteins. The target protein can be purified directly from the SDS-polyacrylamide gel and its sequence can be determined, when the molecular weight of a protein has been revealed.

As a method for isolating proteins binding to the protein of the present invention by using the protein, for example, West-Western blotting analysis (Skolnik, E. Y. et al., Cell (1991) 65, 83-90) can be used. Specifically, a protein binding to the protein of the present invention can be obtained by preparing a cDNA library from cells, tissues, organs (for example, tissues such as ovary, testis, and placenta or cultured cells) expected to express a protein binding to the protein of the present invention by using a phage vector (e.g., ZAP), expressing the protein on LB-agarose, fixing the protein expressed on a filter, reacting the purified and labeled protein of the present invention with the above filter, and detecting the plaques expressing proteins bound to the protein of the present invention according to the label. The protein of the invention may be labeled by utilizing the binding between biotin and avidin, or by utilizing an antibody that specifically binds to the protein of the present invention, or a peptide or polypeptide (for example, GST) that is fused to the protein of the present invention. Methods using radioisotope or fluorescence and such may be also used.

Alternatively, in another embodiment of the screening method of the present invention, a two-hybrid system utilizing cells may be used

("MATCHMAKER Two-Hybrid system", "Mammalian MATCHMAKER Two-Hybrid Assay Kit", "MATCHMAKER one-Hybrid system" (Clontech); "HybriZAP Two-Hybrid Vector System" (Stratagene); the references "Dalton S, and Treisman R (1992) Cell 68, 597-612", "Fields S. and Sternglanz R. Trends Genet. (1994) 10:286-292").

In the two-hybrid system, the protein of the invention is fused to the SRF-binding region or GAL4-binding region and expressed in yeast cells. A cDNA library is prepared from cells expected to express a protein binding to the protein of the invention, such that the library, when expressed, is fused to the VP16 or GAL4 transcriptional activation region. The cDNA library is then introduced into the above yeast cells and the cDNA derived from the library is isolated from the positive clones detected (when a protein binding to the protein of the invention is expressed in yeast cells, the binding of the two activates a reporter gene, making positive clones detectable). A protein encoded by the cDNA can be prepared by introducing the cDNA isolated above to *E.coli* and expressing the protein.

As a reporter gene, for example, Ade2 gene, lacZ gene, CAT gene, luciferase gene and such can be used besides HIS3 gene.

A compound binding to the protein of the present invention can be screened using affinity chromatography. For example, the protein of the invention may be immobilized on a carrier of an affinity column, and a test sample, containing a protein capable of binding to the protein of the invention is supposed to be expressed, is applied to the column. A test sample herein may be, for example, cell extracts, cell lysates, etc. After loading the test sample, the column is washed, and proteins bound to the protein of the invention can be prepared.

The amino acid sequence of the obtained protein is analyzed, an oligo DNA is synthesized based on the sequence, and cDNA libraries are screened using the oligo DNA as a probe to obtain a DNA encoding the protein.

A biosensor using the *surface plasmon resonance phenomenon may* be used as a mean for detecting or quantifying the bound compound in the present invention. When such a biosensor is used, the interaction between the protein of the invention and a test compound can be observed real-time as a surface plasmon resonance signal, using only a minute amount of protein and without labeling (for example, BIAcore, Pharmacia). Therefore, it is possible to evaluate the binding between the protein of the invention and a test compound using a biosensor such as BIAcore.

The methods of screening for molecules that bind when the immobilized

protein of the present invention is exposed to synthetic chemical compounds, or natural substance banks, or a random phage peptide display library, or the methods of screening using high-throughput based on combinatorial chemistry techniques (Wrighton Nc, Farrel FX, Chang R, Kashyap AK, Barbone FP, Mulcahy LS, Johnson DL, Barret RW, Jolliffe LK, Dower WJ; Small peptides as potent mimetics of the protein hormone erythropoietin, Science (UNITED STATES) Jul 26 1996, 273 p458-64, Verdine GL., The combinatorial chemistry of nature. Nature (ENGLAND) Nov 7 1996, 384, p11-13, Hogan JC Jr., Directed combinatorial chemistry. Nature (ENGLAND) Nov 7 1996, 384 p17-9) to isolate not only proteins but chemical compounds that bind to protein of the present invention (including agonist and antagonist) are well known to one skilled in the art.

A compound isolated by the screening is a candidate for drugs which promote or inhibit the activity of the protein of the present invention, for treating or preventing diseases attributed to, for example, cell proliferative diseases such as cancer. A compound in which a part of the structure of the compound obtained by the present screening method having the activity of binding to the protein of the present invention is converted by addition, deletion and/or replacement, is included in the compounds obtained by the screening method of the present invention.

Moreover the present invention provides a method for screening a compound which promotes or inhibits the activity of the protein of the present invention. Since the ZNFN3A1 protein of the present invention has the activity of promoting cell proliferation, a compound which promotes or inhibits this activity of a ZNFN3A1 protein of the present invention can be screened using this activity as an index.

This screening method includes the steps of: (a) culturing cells which express ZNFN3A1 protein in the presence of the subject sample, (b) detecting the proliferation of the cells, and (c) selecting a compound which promotes or inhibits the proliferation in comparison with the proliferation detected in the absence of the subject sample. Compounds that inhibit the expression and/or activity of ZNFN3A1 find utility as anti-cancer agents.

Any ZNFN3A1 proteins can be used for screening so long as they comprise the activity of inhibiting cell proliferation. For example, a human ZNFN3A1 protein can be used and proteins functionally equivalent to these proteins can also be used. ZNFN3A1 proteins may be expressed endogenously or exogenously by cells.

Any subject samples, for example, cell extracts, cell culture supernatant, products of fermenting microorganism, extracts of marine organism, plant extracts, purified or crude proteins, peptides, non-peptide compounds, synthetic micromolecular compounds, natural compounds, can be used. A compound obtained by the above screening for compounds that bind to the protein of the present invention can be also used as the subject compound.

The compound isolated by this screening is a candidate for agonists or antagonists of the protein of the present invention. The term "agonist" refers to molecules that activate the function of the protein of the present invention by binding thereto. Likewise, the term "antagonist" refers to molecules that inhibit the function of the protein of the present invention by binding thereto. Moreover, a compound isolated by this screening is a candidate for compounds which inhibit the *in vivo* interaction of the protein of the present invention with molecules (including DNAs and proteins) .

Cell proliferation can be detected, for example, by determining the speed of cell proliferation, measuring the cell cycle and such, as well as by measuring the colony forming activity as described in the Examples.

The compound isolated by the screening is a candidate for drugs which inhibit the activity of the protein of the present invention and can be applied to the treatment of diseases associated with the protein of the present invention, for example, cancer, more particularly hepatocellular carcinoma.

Moreover, compound in which a part of the structure of the compound inhibiting the activity of ZNFN3A1 proteins is converted by addition, deletion and/or replacement are also included in the compounds obtainable by the screening method of the present invention.

When administrating the compound isolated by the method of the invention as a pharmaceutical for humans and other mammals, such as mice, rats, guinea-pigs, rabbits, chicken, cats, dogs, sheep, pigs, cattle, monkeys, baboons, chimpanzees, the isolated compound can be directly administered or can be formulated into a dosage form using known pharmaceutical preparation methods. For example, according to the need, the drugs can be taken orally, as sugar-coated tablets, capsules, elixirs and microcapsules, or non-orally, in the form of injections of sterile solutions or suspensions with water or any other pharmaceutically acceptable liquid. For example, the compounds can be mixed with pharmacologically acceptable carriers or medium, specifically, sterilized water, physiological saline, plant-oil, emulsifiers, suspending agents, surfactants,

stabilizers, flavoring agents, excipients, vehicles, preservatives, binders and such, in a unit dose form required for generally accepted drug implementation. The amount of active ingredients in these preparations makes a suitable dosage within the indicated range acquirable.

Examples of additives that can be mixed to tablets and capsules are, binders such as gelatin, corn starch, tragacanth gum and arabic gum; excipients such as crystalline cellulose; swelling agents such as corn starch, gelatin and alginic acid; lubricants such as magnesium stearate; sweeteners such as sucrose, lactose or saccharin; flavoring agents such as peppermint, Gaultheria adenothrix oil and cherry. When the unit dosage form is a capsule, a liquid carrier, such as oil, can also be further included in the above ingredients. Sterile composites for injections can be formulated following normal drug implementations using vehicles such as distilled water used for injections.

Physiological saline, glucose, and other isotonic liquids including adjuvants, such as D-sorbitol, D-mannnose, D-mannitol, and sodium chloride, can be used as aqueous solutions for injections. These can be used in conjunction with suitable solubilizers, such as alcohol, specifically ethanol, polyalcohols such as propylene glycol and polyethylene glycol, non-ionic surfactants, such as Polysorbate 80 (TM) and HCO-50.

Sesame oil or Soy-bean oil can be used as a oleaginous liquid and may be used in conjunction with benzyl benzoate or benzyl alcohol as a solubilizers and may be formulated with a buffer, such as phosphate buffer and sodium acetate buffer; a pain-killer, such as procaine hydrochloride; a stabilizer, such as benzyl alcohol, phenol; and an anti-oxidant. The prepared injection may be filled into a suitable ampule.

Methods well known to one skilled in the art may be used to administer the inventive pharmaceutical compound to patients, for example as intraarterial, intravenous, percutaneous injections and also as intranasal, transbronchial, intramuscular or oral administrations. The dosage and method of administration vary according to the body-weight and age of a patient and the administration method; however, one skilled in the art can routinely select them. If said compound is encodable by a DNA, the DNA can be inserted into a vector for gene therapy and the vector administered to perform the therapy. The dosage and method of administration vary according to the body-weight, age, and symptoms of a patient but one skilled in the art can select them suitably.

For example, although there are some differences according to the

symptoms, the dose of a compound that binds with the protein of the present invention and regulates its activity is about 0.1 mg to about 100 mg per day, preferably about 1.0 mg to about 50 mg per day and more preferably about 1.0 mg to about 20 mg per day, when administered orally to a normal adult (weight 60 kg).

When administering parenterally, in the form of an injection to a normal adult (weight 60 kg), although there are some differences according to the patient, target organ, symptoms and method of administration, it is convenient to intravenously inject a dose of about 0.01 mg to about 30 mg per day, preferably about 0.1 to about 20 mg per day and more preferably about 0.1 to about 10 mg per day. Also, in the case of other animals too, it is possible to administer an amount converted to 60kgs of body-weight.

Moreover, the present invention provides a method of diagnosing cancer using ZNFN3A1 gene as a diagnostics marker. This diagnosing method comprises the steps of:

(a) determining a expression level of the ZNFN3A1 gene in biological sample of specimen;

(b) comparing the expression level of ZNFN3A1 gene with that in normal sample, and

(c) defining a high expression level of the ZNFN3A1 gene in the sample as having a cancer.

The expression levels of ZNFN3A1 gene in a particular specimen can be estimated by quantifying mRNA corresponding to or protein encoded by the ZNFN3A1 gene. Quantification methods for mRNA are known to those skilled in the art. For example, the levels of mRNAs corresponding to the ZNFN3A1 gene can be estimated by Northern blotting or RT-PCR. Since all the nucleotide sequences of the ZNFN3A1 gene is shown in SEQ ID NO:1, anyone skilled in the art can design the nucleotide sequences for probes or primers to quantify the ZNFN3A1 gene.

Also the expression level of the ZNFN3A1 gene can be analyzed based on activity or quantity of protein encoded by the ZNFN3A1 gene. A method for determining the quantity of the ZNFN3A1 protein is shown in bellow. For example, immunoasssay method is useful for determination of the protein in biological material. Any biological materials can be used for the determination of the protein or it's activity. For example, blood sample is analyzed for estimation of the protein encoded by serum marker. Another hand, a suitable method can be selected for the determination of the activity protein encoded by the ZNFN3A1 gene according

to the activity of each protein to be analyzed.

Expression levels of the ZNFN3A1 gene in a specimen (test sample) are estimated and compared with those in a nomal sample. When such a comparison shows that the expression level of ZNFN3A1 gene is higher than those in the nomal sample, the subject is judged to be affected with a cancer. The expression level of ZNFN3A1 gene in the specimens from the nomal sample and subject may be determined at the same time. Alternatively, normal ranges of the expression levels can be determined by a statistical method based on the results obtained by analyzing the expression level of ZNFN3A1 gene in specimens previously collected from a control group. A result obtained by examining the sample of a subject is compared with the normal range; when the result does not fall within the normal range, the subject is judged to be affected with the cancer. In the present invention, the cancer to be diagnosed is preferably a hepatocellular carcinoma.

In the present invention, a diagnostic agent for diagnosing hepatocellular carcinoma is also provided. The diagnostic agent of the present invention comprising a compound that binds to the DNA or the protein of the present invention. Preferably, the oligonucleotide that hybridize to polynucleotide of the present invention, or the antibodies may that bind to the protein of the present invention may be used as these compound.

The following examples are presented to illustrate the present invention and to assist one of ordinary skill in making and using the same. The examples are not intended in any way to otherwise limit the scope of the invention.

Any patents, patent applications, and publications cited herein are incorporated by reference.

EXAMPLES

Best Mode for Carrying out the Invention

The present invention is illustrated in details by following Examples, but are not restricted to these Examples.

Materials and Methods

A variety of mammalian host cells including human hepatoma cell line Huh7 and Alexander, human cervix cell line HeLa and mouse fibroblast cell line NIH3T3, all of which were obtained from the American Type Culture Collection (ATCC) were used to isolate and characterize *ZNFN3A1*. Human hepatoma cell lines SNU423 and SNU475 were also used and were obtained from the Korea cell-line

bank. All cell lines were grown in monolayers in appropriate media: Dulbecco's modified Eagle's medium for Huh7 Alexander and NIH3T3; Eagle's Minimum Essential Medium for HeLa; RPM1640 for SNU423 and SNU475 supplemented with 10% fetal bovine serum and 1% antibiotic/antimycotic solution (Sigma) and grown at 37°C in air containing 5% CO_2.

Cloning of *ZNFN3A1* was generally done by PCR using KOD-plus (TOYOBO). For E. Coli expression, coding region of *ZNFN3A1* was cloned in the *Eco*R I-*Kpn* I site of pET21a.

For mammalian cell expression, coding region of *ZNFN3A1* was cloned in the *Eco*R I-K*pn* I site of pcDNA3.1 (+) and (-) (Invitrogen), *Eco*R I-*Kpn* I site ot pFLAG and *Eco*R I-*Kpn* I site of pEGFP (Clontech). Coding region of *KIAA0054* was cloned in the *Eco*R I-*Xho* I site of pCMV-HA (Clontech).

The RNA was prepared in the following experiments by extracting total RNA with the Qiagen RNeasy kit (Qiagen) or Trizol reagent (Life Technologies) according to the manufacturer's protocol.

The RNA was amplified herein by RT-PCR. 10 µg of total RNA was reversely transcribed for single-stranded cDNAs using poly $dT_{12,18}$ primer (Amersham Biosciences) with Superscript II reverse transcriptase (Life Technologies). Each single-stranded cDNA was diluted for subsequent PCR amplification. Standard RT-PCR was carried out in a 20 µl volume of PCR buffer (TAKARA), and amplified for 4 min at 94°C for denaturing, followed by 20 (for *GAPDH*) or 30 (for *ZNFN3A1*) cycles of 94°C for 30 s, 56°C for 30 s, 72°C for 30 s, in the Gene Amp PCR system 9700 (Perkin-Elmer). Primer sequence were as follows, for GAPDH forward; 5'-ACAACAGCCTCAAGATCATCAG (SEQ. ID. NO. 4) and reverse; 5'- GGTCCACCACTGACACGTTG(SEQ. ID. NO. 5), for *ZNFN3A1* forward; 5'-TTCCCGATATCAACATCTACCAG(SEQ. ID. NO. 6) and reverse; 5'-AGTGTGTGACCTCAATAAGGCAT(SEQ. ID. NO. 7).

For detection of ZNFN3A1, rabbit anti-ZNFN3A polyclonal antibody was generated. Full coding sequence of *ZNFN3A1* was amplified by PCR reaction using testis cDNA as a template and cloned in pET21 a (Novagen). The cloned vector was transfected into BL21-CodonPlus® competent cells (Stratagene). Recombinant ZNFN3A1 protein was induced by 1.0 mM IPTG at 30°C for 6 h. His-ZNFN3A1 fusion protein was purified using Pro Bond™ Resin (Invitrogen). Rabbits were immunized ten times with purified His-ZNFN3A1. Immunoblotting with this polyclonal antibody showed single 50 kD band of FLAG-tagged ZNFN3A1, which was identical pattern to that detected using anti-FLAG monoclonal antibody

(Sigma) (data not shown).

The effect of ZNFN3A1 in cell cycle progression was determined by flow cytometry. Cells were plated at a density of $1X10^5$ cells/100 mm dish. The cells were trypsinized at the given time course, collected in PBS and fixed in 70% cold ethanol. After RNase treatment, cells were stained with propidium iodide (50 µg/ml) in PBS. Flow cytometry was performed on a Becton Dickinson FACScan and analyzed by CellQuest and ModFit software (Verity Software House), The percentages of nuclei in G0/G1, S and G2/M phases of the cell cycle, and any sub-G1 population were determined from at least 20,000 ungated cells.

Immunoblotting and immunohistochemistry was also performed in various experiments. For immunoblotting, the cells were washed twice with PBS and harvested in lysis buffer (150 mM NaCl, 1% Triton X-100, 50mM Tris-HCl pH 7.4, 1mM DTT, and 1X complete Protease Inhibitor Cocktail (Roche)). After the cells were homogenized and centrifuged at 10,000*xg* for 30 min, the supernatant were standardized for protein concentration by the Bradford assay (Bio-Rad). Proteins were separated by 10% SDS-PAGE and Immunoblotted with mouse anti-FLAG, rabbit anti-RNA polymerase II, rabbit anti-HA antibody and rabbit anti-ZNFN3A1 antibody. HRP-conjugated goat anti-mouse IgG and anti-rabbit IgG served as the secondary antibody for the ECL Detection System (Amersham Biosciences).

For immunohistochemistry, immunohistochemical staining was carried out using anti-ZNFN3A1 antibody. Paraffin-embedded tissue sections were subjected to the SAB-PO peroxidase immunostaining system (Nichirei, Tokyo, Japan) according to the manufacturer's recommended method. Antigens were retrieved from deparaffinized and re-hydrated tissues by pretreating the slides in citrate buffer (pH6) in a microwave oven for 10 mm at 700W.

<u>Example 1: Identification of a novel gene frequently up-regulated in HCCs</u>

Using genome-wide cDNA microarray with 23040 genes, we identified an expressed sequence tag (EST), which was commonly up-regulated in hepatitis B-positive and/or hepatitis C-positive HCCs. Among them, we focused on a gene, *A6681*, corresponding to an EST (Hs. 8109), because its expression was significantly up-regulated in eleven of twelve (91.7%) clinical HCCs compared with the corresponding noncancerous liver tissues (Figure 1A). The elevated expression of the gene *A6681* was also confirmed in another 10 HCC cases by RT-PCR (Figure 1B). The relative expression confirmed by semi-quantitative RT-PCR were well correlated to those confirmed by cDNA microarray.

<u>Example 2A: Isolation and characterization of a novel human gene, *ZNFN3A1*</u>

Multi-tissue northern blots from Clonetech were used to hybridize a ^{32}P labeled partial *A6681* cDNA with an approximately 1.7-kb transcript expressed in testis and skeletal muscle purchased from Clontech. (Figure 1C) The A6681 probe was prepared by RT-PCR using a set of primers, 5'-TTCCCGATATCAACATCTACCAG-3' (SEQ. ID. NO. 6) and 5'-AGTGTGTGACCTCAATAAGGCAT-3'(SEQ. ID. NO. 7) . Pre-hybridization, hybridization and washing were performed according to the supplier's recommendations. The blots were autoradiographed with intensifying screens at -80°C for 24 h. The approximately 1.7 kb polynucleotide transcript was termed the *ZNFN3A1* gene.

The sequence of the 5' region of the transcript was determined by 5' rapid amplification of cDNA ends (5'RACE) which was performed using the Marathon cDNA amplification kit by Clontech in accordance with the manufacturer's instructions. For the amplification of the 5' part of *ZNFN3A1* cDNA, a gene specific reverse primer (5'-CTGCCAAGAAGTCGGAGTCTGGAG) [SEQ. ID. NO. 8] was used. The cDNA template was synthesized from human testis mRNA by RT-PCR. The PCR products were cloned using TA cloning kit by Invitrogen and their sequences were determined with an ABI PRISM 377 DNA sequencer from Applied Biosystems. As a result, 1622 nucleotides sequence was assembled and obtained as shown and described in SEQ. ID. NO: 1 that contained an open reading frame of 1284 nucleotides encoding 428 amino acids.

Because no EST clones containing the 5' part of the *ZNFN3A1* gene was identified in EST data bases: Genomic sequences were searched corresponding to the ZNFN3A1 cDNA in the genomic databases. Cosmid sequences were found in the genomic sequence comparison that were assigned to chromosomal band 1q44, which also included the *ZNFN3A1* gene. Using GENSCAN and Gene Recognition (GENSCAN from MIT (http://genes.mit.edu/GENSCAN.html) and GRAIL 2 from IMS (http://www.genome.ad.jp)) and Assembly Internet Link program (AutoAssembler 2.1) provided by ABI, the candidate-exon sequences were predicted and exon-connection was performed.

Example 2B: Isolation and characterization of a novel human protein, ZNFN3A1

With the 1622 nucleotide sequence that contained an open reading frame encoding 428 amino acids, the protein ZNFN3A1 was identified using a Simple Modular Architecture Research Tool (SMART) from EMBL (http://smart.embl-heidelberg.de). SMART suggested that the ZNFN3A1 protein contained a zf-MYND [zinc finger protein (MYND domain containing)] domain (codons 49-87) as

well as a SET [(Su (var) 3-9, Enhancer-of-zeste, Trithorax)] domain (codons 117-246) (Figure 2A). The amino acid sequence of the ZNFN3A1 protein shared 94% identity with amino acid sequence of the Mus musculus ES cells cDNA (GenBank Accession number: AK010447) (Figure 2B) and the gene encoding the protein was termed "ZNFN3A1"/ (zinc finger protein, subfamily 3A (MYND domain containing), 1) by nomenclature committee.

Example 3: Sub-cellular localization of ZNFN3A1

The entire coding region corresponding to ZNFN3A1 was cloned into a pEGFP-N1 vector and a pFLAG-CMV-5a vector, and these constructs were transfected into SNU475 cells and expressed. Expression of EGFP tagged ZNFN3A1 and FLAG tagged ZNFN3A1 was confirmed by western blotting (data not shown). Both ZNFN3A1-EGFP fusion protein and FLAG-tagged ZNFN3A1 protein were detected homogenously in cytoplasm and nucleus by fluorescent immunocytochemistry (Figure 3A-F). Sub-cellular localization of endogenous ZNFN3A1 protein was observed by using specific antibody against ZNFN3A1.

Interestingly subcellular localization of ZNFN3A1 protein is altered during cell cycle progression or due to the density of cultured cells. Performing immunocytochemistry for endogenous ZNFN3A1 protein, some amount of proteins were localized in nucleus at the case of low cell concentration culture (Figure 3G-I and Figure 3M-O). However, in the case of high cell concentration culture, most of ZNFN3A1 protein localized in cytoplasm (Figure 3J-L and Figure 3P-R). These results revealed that sub-localization of ZNFN3A1 depends on cell concentration. ZNFN3A1 accumulates in the nucleus when cells are in middle to late S phase or cultured in sparse condition, while ZNFN3A1 localizes in the cytoplasm as well as nucleus when they are in other phases or grown in dense condition.

The immunocytochemistry was performed by fixing cultured cells on chamber slides with PBS containing 4% paraformaldehyde for 15 min, then rendered permeable with PBS containing 0.1% Triton X-100 for 2.5 min at RT. The cells were covered with 2% BSA in PBS for 24 h at 4°C to block non-specific hybridization and subsequently incubated with mouse anti-FLAG antibody at 1:2000 dilution and rabbit anti-ZNFN3A1 antibody at 1:3000 for the first antibody. Antibodies were stained fluorescent substrate conjugated anti-mouse IgG and anti-rabbit IgG second antibody (ICN/Cappel and Jackson Immuno Research). Nuclei were counter stained by 4', 6'-diamidine-2'-phenylindole dihydrochioride (DAPI). Fluorescence image was obtained with an ECLIPSE E800 microscope.

Localization of ZNFN3A1 may be dependant on cell cycle and thus cell

cycle was analyzed using flow cytometry in different cell concentration of SNU475 cells. Cells were plated at a density of $1X10^5$ cells/100 mm dish. The cells were trypsinized at the given time course, collected in PBS and fixed in 70% cold ethanol. After RNase treatment, cells were stained with propidium iodide (50 µg/ml) in PBS. Flow cytometry was performed on a Becton Dickinson FACScan and analyzed by CellQuest and ModFit software (Verity Software House). The percentages of nuclei in G0/G1, S and G2/M phases of the cell cycle, and any sub-G1 population were determined from at least 20,000 ungated cells.

Compared with low cell concentration and high cell concentration, the population of cells in G0/G1 phase increased in the case of high cell concentration, however the population of cells in S and G2/M phase decreased drastically. To determine the effect of cell cycle to localization of ZNFN3A1 in detail, Huh7 cells were synchronized using aphidicolin and sub-cellular localization of ZNFN3A1 was observed (Figure 4a, b). Most of the Huh7 cells stayed G0/G1 phase 36 h after treatment of aphidicolin and the ZNFN3A1 was localized in the cytoplasm. When aphidicolin was removed from culture medium, cell cycle was progressed and ZNFN3A1 protein moved to nucleus from cytoplasm. These data showed that sub-cellular localization of ZNFN3A1 protein was regulated by cell cycle status, and ZNFN3A1 protein moved to nucleus in the proliferative condition.

<u>Example 4: Promotion of growth of normal tissue cell line NIH3T3 cells by ZNFN3A1</u>

To analyze the effects of *ZNFN3A1* gene transfer on growth of hepatoma cell lines, normal tissue cell line NIH3T3 cells were transfected with an expression plasmid pcDNA3.1 containing sense *ZNFN3A1* and anitsense-*ZNFN3A1*.

NIH3T3 cells were transfected by pcDNA3.1 vectors (Invitrogen) containing full coding sequence of *ZNFN3A1* using FuGENE 6 transfection reagent according to the supplier's recommendations. Cells were maintained in DMEM containing 10% FBS and 0.9 mg/ml geneticin, and single colonies were selected. Constitutive *ZNFN3A1* expression was determined by RT-PCR.

The NIH3T3 cells generally showed no endogenous expression of *ZNFN3A1* mRNA. Upon colony formation, the sense *ZNFN3A1* expression vector promoted colony formation in NIH3T3 cells compared with mock and antisense-*ZNFN3A1* vectors demonstrated no growth as shown in Figures 5A, 5B.

Colony formation assays were performed on the plating cells at a density of $1X10^5$ cells/100mm dish. After 24 h, the cells were transfected by plasmid vector using FuGENE 6 transfection reagent (Roche) and were cultured with appropriate

concentration of geneticin for 2 weeks. Cells were fixed with 100% methanol and stained by Giemsa solution. These colony formation assays were confirmed by three independent experiments.

To further investigate the growth promotive effects of *ZNFN3A1*, stable NIH3T3 transfectant cells of *ZNFN3A1* were further examined. The sense *ZNFN3A1* stable transfectant cells expressed constitutive *ZNFN3A1* mRNA (Figure 5C, 5D). Expression was determined by RT-PCR. As shown Figure 5C, 5D, the clone which expressed *ZNFN3A1* constantly showed high growth ability compared with antisense ZNFN3A1 and control vector transfectant cells showing that ZNFN3A1 plays an important role for growth promotion of hepatocellular carcinoma cells.

Example 5: Reduced expression of *ZNFN3A1* by antisense oligonucleotides suppresses growth of hepatoma cells

To examine whether suppression of *ZNFN3A1* may induce growth retardation and/or cell death to HCC cells, various antisense S-oligonucleotides were synthesized to suppress *ZNFN3A1* expression. Sense or antisense S-oligonucleotides of *ZNFN3A1* encompassing initiation codon was transfected using LIPOFECTIN Reagent (GIBCO-BRL). The sequences of phosphorothioate-modified ODNs were as follows: antisense S-oligonucleotides; 5'-GCGGGAGGAT GGAGCC (SEQ. ID. NO.3).

The cells were cultured with the antisense and sense S-oligonucleotides for 24 hours, and analyzed for their expression of *ZNFN3A1* and ZNFN3A1 by RT-PCR and western blotting using anti-ZNFN3A1 antibody.

The cells transfected with antisense S-oligonucleotides, encompassing the initiation codon significantly decreased endogenous expression of *ZNFN3A1* in SNU475 and Huh7 cells that constitutively express abundant amount of *ZNFN3A1* (Figure 6A, B). Transfection of antisense S-oligonucleotides also significantly suppressed cell numbers of the Huh7 and SNU475 cells as determined by a colony formation assay as previously described.

An MTT assay was performed by plating cells at a density of $5X10^5$ cells/100 mm dish. On the next day, the cells were transfected in triplicate with sense or antisense S-oligonucleotides of *ZNFN3A1*. After 72 h of culture, the medium was replaced with 10 ml of fresh medium containing 5 mg 3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyltetrazolium in bromide (MTT) (SIGMA). After further 4 h of incubation at 37°C, cells were lysed by the addition of 1 ml of 0.01 N HCl/10%SDS. The color reaction was quantified with an ELISA plate reader at a

test wavelength of 570 nm (reference, 630 nm). The cell viability was represented by the absorbance compared to the control.

Suppression of cell growth from transfection of antisense-*ZNFN3A1* was also seen in other HCC cell lines including Alexander cells and SNU423 cells, both of which also constitutively express abundant amount of *ZNFN3A1* compared with that of control S-oligonucleotides as shown in Figure 6C, 6D. The results of the colony formation assay was confirmed by three independent experiments. Furthermore, flow-cytometry demonstrated that inhibition of *ZNFN3A1* expression significantry decreased numbers of cells in S phase population and increased numbers of sub-G1 phase (Figure 6E). These results revealed that the suppression of *ZNFN3A1* expression induced inhibition of growth and promotion of apoptosis of HCC.

Example 6: Interaction of ZNFN3A1 with RNA helicase KIAA0054

To examine the oncogenic mechanism of *ZNFN3A1*, ZNFN3A1-interacting proteins were searched using yeast two-hybrid screening system. Yeast two-hybrid assay was performed with MATCHMAKER GAL4 Two-Hybrid System 2 and System 3 by Clontech according to the manufacturer's protocols. The full coding sequence of *ZNFN3A1* was cloned in *Eco*R I site of pAS2-1 as a bait vector. For library screening, a human testis library was cloned in pACT2 (Clontech). Simultaneously co-transformed AH109 yeast cells (with pAS2-1 bait vector and a variety of pACT2 prey vectors) were plated on SD minimal Medium (-Ade/-His/-Leu/-Trp) in addition of 25 mg/L X-α-Gal (Clontech) and 25 mM 3-amino-1,2,4,-trizole (Sigma). Library plasmid was isolated from the positive colonies and the sequence and the frame was confirmed.

Among the clones identified, the C-terminal region of RNA helicase KIAA0054 interacted with ZNFN3A1 by simultaneous transformation using pAS2.1-ZNFN3A1 and pACT2-KIAA0054 (Figure 7A, 7B). RNA helicases constitute a family of proteins that unwind double-stranded RNA by using nucleoside triphosphates as a source of energy. It is clear that RNA helicases are a widely dispersed group of proteins found in virtually all biological processes.

To further confirm the interaction of ZNFN3A1 with KIAA0054 , FLAG-tagged ZNFN3A1 protein was prepared. HeLa cells were transfected with 8 μg pFLAG-CMV-ZNFN3A1 and pCMV-HA-KIAA0054 DNA per 10-cm dish and collected after an additional 48 h. Cells were washed once with 10 mM sodium phosphate buffer (pH 7.4) containing 150 mM NaCl (PBS) and lysed in NET-N buffer (150 mM NaCl, 0.5% NP-40, 20mM Tris-HCl pH8.0, 1 mM EDTA, and 1X

complete Protease Inhibitor Cocktail). In a typical immunoprecipitation reaction, 300 µg of the HeLa whole-cell lysate extract was incubated with 1µg of the desired antibody and 20 µl of protein A or protein G Sepharose beads (Zymed) at 4°C for 1-2 hr. Beads were washed four times in 1 ml of NET-N buffer. Proteins bound to the beads were eluted by boiling in SDS sample buffer, separated by SDS-PAGE, and transferred to nitrocellulose membranes. The membranes were used for immunoblotting as described above. The lysate was directly analyzed by immunoblotting with anti-FLAG antibody and anti-HA antibody as a control. The ZNFN3A1 protein revealed its association to HA-tagged KIAA0054 protein expressed in HeLa mammalian cells (Figure 7C).

To identify the region of ZNFN3A1 responsible for its interaction with KIAA0054, two-hybrid assay was performed using deletion fragments of ZNFN3A1. Mutants lacking amino acids 1 to 250 or 100 to 428 were negative for interaction with KIAA0054 C-terminal region, whereas a fragment containing amino acids 100 to 250 was positive (Figure 8). This indicates that the ZNFN3A1-binding region resides in the region of the SET domain.

Example 7: Interaction among ZNFN3A1, RNA helicase and RNA polymerase II

RNA helicase plays a crucial role for transcription by binding to transcription factor and RNA polymerase II. These results revealed that there was a possibility that zinc finger protein ZNFN3A1 regulated transcription through association with not only RNA helicase KIAA0054 but also RNA polymerase II. The association among ZNFN3A1, RNA helicase KIAA0054 and RNA polymerase II was tested by co-immunoprecipitation (Figure 9). HeLa cells were transfected with 8 µg pFLAG-CMV-ZNFN3A1 and pCMV-HA-KIAA0054 DNA per 10-cm dish and collected after an additional 48 h. Cells were washed once with 10 mM sodium phosphate buffer (pH 7.4) containing 150 mM NaCl (PBS) and lysed in NET-N buffer (150 mM NaCl, 0.5% NP-40, 20mM Tris-HCl pH8.0, 1 mM EDTA, and 1X complete Protease Inhibitor Cocktail). In a typical immunoprecipitation reaction, 300 µg of whole-cell extract was incubated with 1µg of antibody and 20 µl of protein A or protein G Sepharose beads (Zymed) at 4°C for 1-2 hr. Beads were washed four times in 1 ml of NET-N buffer. Proteins bound to the beads were eluted by boiling in SDS sample buffer, separated by SDS-PAGE, and transferred to nitrocellulose membranes. The membranes were used for immunoblotting as described above.

An association with endogenous RNA polymerase II was detected by using anti-FLAG antibody, anti-HA antibody and anti-RNA polymerase II antibody. When an anti-FLAG antibody was used for immunoprecipitation and an anti-RNA

polymerase II antibody was used for immunoblotting, RNA polymerase II specific band was detected strongly in the case of expressing FLAG-ZNFN3A1 protein and HA-KIAA0054 protein together, and detected marginally in the case of expressing only FLAG-ZNFN3A1 protein. On the other hand, when anti-HA antibody was used for immunoprecipitation and anti-RNA polymerase II antibody was used for immunoblotting, RNA polymerase specific band was detected strongly not only in the case of expressing FLAG-ZNFN3A1 protein and HA-KIAA0054 protein together, but also expressing HA-KIAA0054 alone. Moreover, co-immunoprecipitation was worked reversely by changing antibody for immunoprecipitation and immunoblotting. The results were similar in the case of using opposite antibody. These results show that RNA helicase KIAA0054 can mediate complex formation between ZNFN3A1 protein and RNA polymerase II via contacts with each protein. Thus, transcriptional regulation by ZNFN3A1, RNA helicase KIAA0054 and RNA polymerase II may play an important role for promotion of cell growth in hepatocellular carcinogenesis. Figure 12 is an illustration of complex formed between ZNFN3A1, RNA helicase, and RNA polymerase II to regulate transcription of gene.

Example 8: Sequence-specific binding of ZNFN3A1 protein

A consensus DNA binding site for ZNFN3A1 was determined by using an oligonucleotide construct containing a core of 20 random nucleotides. The putative consensus sequences were searched that were able to associate with ZNFN3A1 protein in vitro by means of DNA selection using random oligonucleotides. A GST–ZNFN3A1 fusion protein was prepared and immobilized with Sepharose 4B, and double-stranded random oligonucleotide DNAs were selected that associated with the protein. After ten times of selection and subsequent amplification, the amplified DNA was cloned into pCR vector (Clontech) and 92 clones were sequenced. Among the 92 clones, 32 (34.8%) contained a consensus sequence of 5'-CCCTCC-3', which is 102-fold higher incidence than calculated probability (Figure 10A). To prepare recombinant protein expressing the zinc finger domain of ZNFN3A1, a cDNA fragment corresponding codons full length of *ZNFN3A1* was amplified by RT-PCR using a set of primers, 5'-CGGAATTCATGGAGCCGCTGAAGGTGGAAAAG-3' [SEQ. ID. NO. 9] and 5'-CCGCTCGAGGGATGCTCTGATGTTGGCGTCG-3' [SEQ. ID. NO. 10], which was subcloned into an appropriate cloning site of pGEX-6P plasmid. Recombinant fusion protein was prepared and purified by Sepharose 46 column as described previously. Oligonucleotides with the sequence "5'-

GGGAGAATTCCGACACGCGT(N20)CTCGAGCGTCTACATGGATCCTCA-3'" [SEQ. ID. NO. 11], were used for selection and amplification as previously described.

By using Eukaryotic Promoter Database (http://www.epd.isb-slb.ch/index.html), candidate downstream genes of ZNFN3A1 was searched and five candidate genes were picked. First, the expression level of each genes was determined by semi-quantitative RT-PCR in COS7 stable transformant that expressed ZNFN3A1 exogenousiy (Figure 10B). These results revealed that the expression of EGFR, c-myc and Bc12 was up-regulated by ZNFN3A1.

Four putative binding sites for ZNFN3A1, whose consensus target sequence is 5'-(C)CCCTCC(T) or (A)GGAGGG(G)-3', were identified in the 5' flanking region of EGFR, between -213-bp and -207-bp (CBS1), between -106-bp and -100-bp (CBS2), between -65-bp and -59-bp (CBS3) and between -46-bp and -40-bp (CBS4). A wild-type reporter plasmid (P1) containing CBS1 CBS2, cBS3, and CBS4 as well as four deletion constructs of the plasmid (P2, P3, P4 and P5) were prepared by cloning each sequence into the appropriate enzyme sites of pGL3-Basic vector (Promega). Plasmids P1, P2, P3, P4 and P5 were constructed by amplification of
P1-F (5'-GGGGTACCCAGTGCTGGGAACGCCCCTCTCG-3') [SEQ. ID. NO. 12],
P2-F (5'-GGGGTACCCACTCCCGCCGGAGACTAGGTCC-3') [SEQ. ID. NO. 13],
3-F (5'-GGGGTACCCTCGCATTCTCCTCCTCCTCTGC-3') [SEQ. ID. NO. 14],
4-F (5'- GGGGTACCTGGTCCCTCCTCCTCCCGCCCTG-3') [SEQ. ID. NO. 15] or
P5-F (5'- GGGGTACCTCCCGCCCTGCCTCCCGCGCCTC-3') [SEQ. ID. NO. 16]
with the same reverse primer (P1-R; 5'-GAAGATCTAG GTGGCCTGTC GTCCGGTCTG G-3') [SEQ. ID. NO. 17]. Site-directed mutagenesis was carried out for both putative ZNFN3A1 binding sites, replacing CCCTCC by CATTCC using the QuickChange Site-Directed Mutagenesis Kit according to the supplier's *recommendations (Stratagene).*

Each reporter plasmid (2 μg) was co-transfected with 0.2 μg of pRL-TK plasmid (Promega) using FuGENE6 Reagent (Boehringer) according to the manufacturer's instructions. Luciferase assays were carried out using a Dual-Luciferase Reporter Assay System (Promega) following the manufacturer's protocol.

<u>Example 9: Up-regulation of *EGFR* promoter activity by ZNFN3A1</u>

Since enhanced expression of EGFR had been reported in HCCs, we focused on that receptor molecule and tested whether its promoter was regulated

by the ZNFN3A1-binding sequences. We identified four possible ZNFN3A1-binding motifs (CBS1, 2, 3 and 4) in its 5' flanking region and prepared reporter plasmids containing the four motifs (P1) as well as various deletion forms (P2, P3, P4 and P5). When these reporter plasmids were transfected into SNU475 cells, the activity of P1 was significantly higher than that of P2, P3, P4 or P5. Given the fact that the activity of P4 was very similar to that of P5, we suspected that a region between –261 and –50, containing CBS1, CBS2 and CBS3, might be associated with transcriptional activation of *EGFR*. To disclose the roles of these domains, we constructed reporter plasmids containing mutant CBS1 (P1m1) and mutant CBS2 (P1m2) or mutant CBS3 (P1m3) in which each of the candidate binding motifs was changed from 5'-CCCTCC-3' to 5'-CATTCC-3' (Figure 11A). Reporter assays revealed that fragments containing mutated motifs (P1m1, P1m2 and P1m3) activated transcription of *EGFR* much more weakly than did P1 (Figure 11B). These results implied that the three putative ZNFN3A1-binding motifs were involved in transcriptional activation of *EGFR*.

Industrial Applicability

The expression of novel human gene *ZNFN3A1* is markedly elevated in hepatocellular carcinoma as compared to non-cancerous liver tissues. Accordingly, this gene may serve as a diagnostic marker of HCC and the protein encoded thereby may be used in diagnostic assays therefore.

The present inventors have also shown that the expression of novel protein ZNFN3A1 promotes cell growth whereas cell growth is suppressed by antisense oligonucleotides corresponding to the ZNFN3A1. These findings suggest that ZNFN3A1 stimulates oncogenic activity. Thus, this novel oncoprotein is a useful target for the development of anti-cancer pharmaceuticals. For example, agents that block the expression of ZNFN3A1 or prevent its activity may find therapeutic utility as anti-cancer agents, particularly anti-cancer agents for the treatment of HCC. Examples of such agents include antisense oligonucleotides and antibodies that recognize ZNFN3A1.

Furthermore, the present inventors have shown that ZNFN3A1 directly associates with an RNA helicase and forms a complex with RNA polymerase II. This complex then activates transcription of downstream target genes, including EGFR, through direct binding of the complex with an element "(C)CCCTCC(T)" in the 5' flanking region. Thus, agents that inhibit the activity of the complex may also find utility in the treatment and prevention of HCC.

While the invention has been described in detail and with reference to specific embodiments thereof, it will be apparent to one skilled in the art that various changes and modifications can be made therein without departing from the spirit and scope of the invention.

CLAIMS

1. An isolated DNA of any one of the following (a) to (d):
   (a) a DNA encoding the protein consisting of the amino acid sequence of SEQ. ID. NO. 2;
   (b) a DNA containing the coding region of the nucleotide sequence of SEQ. ID. NO. 1;
   (c) a DNA encoding a protein consisting of the amino acid sequence of SEQ. ID. NO. 2, in which one or more amino acids are replaced, deleted, inserted, and/or added, such that the encoded protein is a functional equivalent of the protein consisting of the amino acid sequence of SEQ. ID. NO. 2; and
   (d) a DNA hybridizing under stringent conditions with a DNA consisting of the nucleotide sequence of SEQ. ID. NO.1, such that the encoded protein is a functional equivalent to the protein consisting of the amino acid sequence of any one of SEQ. ID. NO.2
2. An isolated DNA encoding a partial peptide of the protein consisting of the amino acid sequence of SEQ. ID. NO.2.
3. A vector into which the DNA of claims 1 or 2 is inserted.
4. A transformed cell harboring the DNA of claims 1 or 2, or the vector of claim 3.
5. An isolated protein encoded by the DNA according to claim 1 having the activity of promoting cell proliferation and transcriptional activation of target genes.
6. An isolated polypeptide encoded by the DNA according to claim 2.
7. A method for producing the protein of claim 5, comprising the steps of culturing the transformed cell of claim 4, and collecting the protein expressed from the cells or the culture supernatant thereof.
8. A transcription activation complex comprising the protein of claim 5 and at least one-co-activator thereof.
9. The complex of claim 8 the co-activator is selected from the group consisting of an RNA helicase, and an RNA polymerase II.
10. An antibody that recognizes and specifically binds to the protein of claim 5 or the complex of claim 8.
11. A polynucleotide hybridizing with a DNA consisting of the nucleotide sequence of SEQ. ID. NO.1 or the complementary strand thereof, and comprising at least 15 nucleotides.
12. A therapeutic agent for treating cancer comprising the antibody of claim 10 or

the polynucleotide of claim 11.

13. A method of screening for a compound that binds to the protein of claim 5, comprising the steps of:
    (a) contacting a subject sample, containing at least one test compound, with the protein of claim 5 or a partial peptide thereof;
    (b) detecting the binding activity of the subject sample with the protein or the partial peptide thereof; and
    (c) selecting the test compound that binds to the protein or the partial peptide thereof.

14. A compound binding to the protein of claim 5, which can be isolated by the method of claim 13.

15. A method of screening for a compound that inhibits the activity of the protein of claim 5, comprising the steps of:
    (a) culturing cells which express the protein of claim 5 or a partial peptide thereof in the presence of a subject sample which contains at least one test compound;
    (b) detecting the proliferation of the cell; and
    (c) selecting the test compound that inhibits the proliferation as compared to the proliferation detected in the absence of the subject sample.

16. A compound that inhibits the activity of the protein of claim 5, which can be isolated by the method of claim 15.

17. A method of screening a compound for anti-cancer activity, comprising the steps of:
    (a) contacting a subject sample, containing at least one test compound, with the protein of claim 5, a co-activator thereof and a DNA containing the target sequence of said protein under suitable conditions to allow formation of the complex of said protein with the DNA; and
    (b) selecting the test compound that inhibits the formation of the complex.

18. The method of claim 17, wherein said target sequence comprises a CBS sequence flanking the 5' region of EGFR.

19. A method of screening a compound for anti-cancer activity, comprising the steps of:
    (a) contacting a subject sample, containing at least one test compound, with the complex of claim 7 and a reporter gene with a transcriptional regulatory region recognized by said complex; and
    (b) selecting the test compound that inhibits the expression of the reporter

gene.

20. An anti-cancer composition comprising a compound isolated by the method of claims 18 or 19 as an active ingredient.

21. An anti-cancer composition comprising an antisense oligonucleotide, ribozyme, or interference RNA that binds to the DNA of claim 1 as an active ingredient.

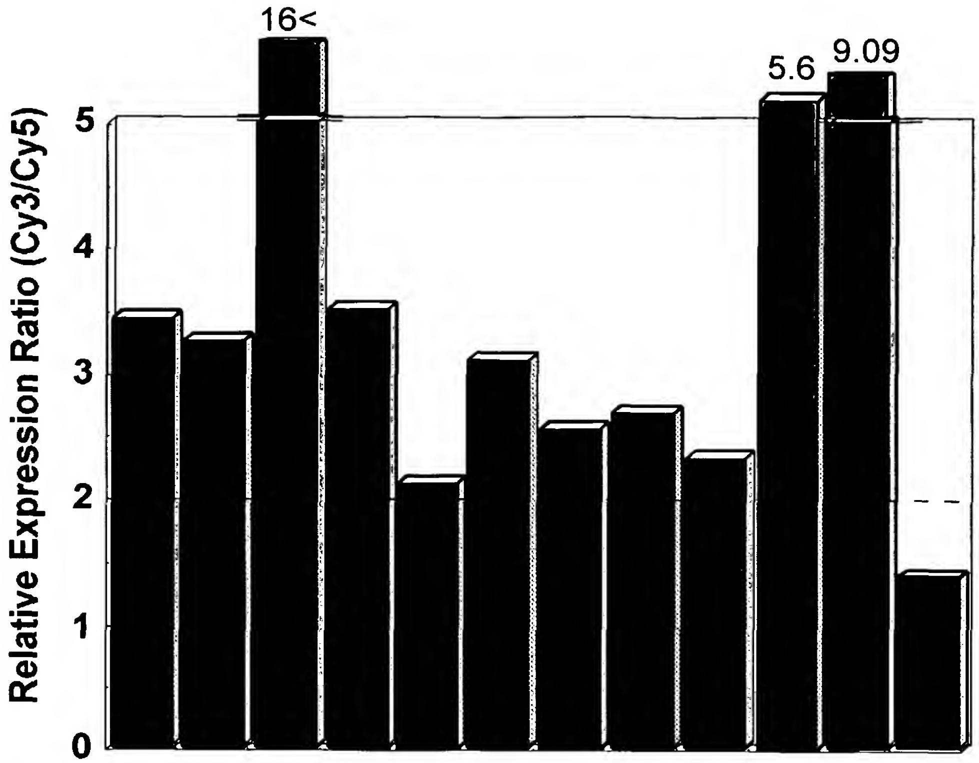
22. A method diagnosing cancer, wherein said method comprises the steps of:
   (a) determining a expression level of the ZNFN3A1 gene in biological sample of specimen;
   (b) comparing the expression level of ZNFN3A1 gene with that in normal sample, and
   (c) defining a high expression level of the ZNFN3A1 gene in the sample as having a cancer.

23. The method of claim 13, wherein the cancer is hepatocellular carcinoma.

24. A diagnostic agent for diagnosing hepatocellular carcinoma comprising a compound that binds to the DNA of claim 1 or the protein of claim 5.

FIG.1

a



Case No.	40	53	90	143	145	148	153	154	156	159	161	177
Ratio	3.4	3.3	16	3.5	2.1	3.1	2.6	2.7	2.3	5.6	9.1	1.4

b

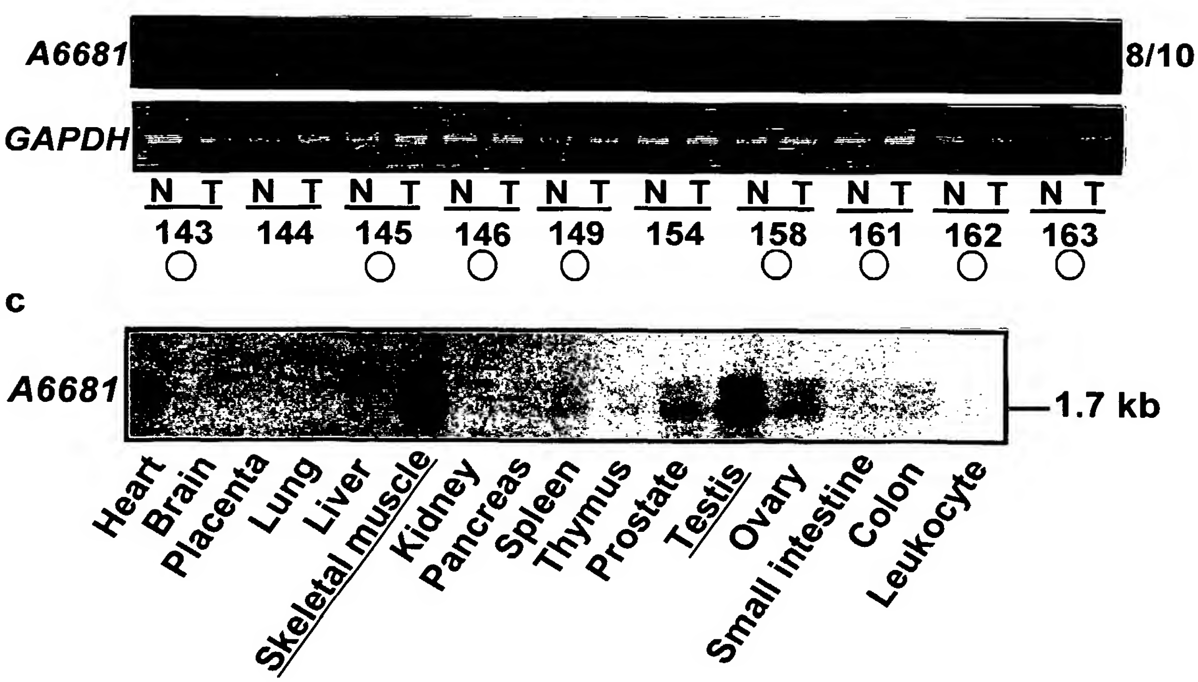
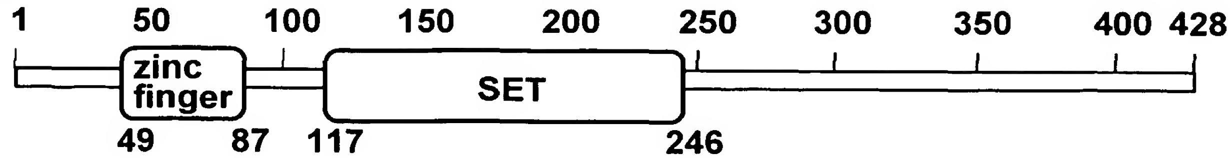


FIG.2

a



b

```
                                                                                 A
ZNFN3A1 : 1    MEPLKVEKFATANRGNGLRAVTPLRPGELLFRSDPLAYTVCKGSRGVVCDRCLLGKEKLM 60
               ME LKVEKF TANRGNGLRAV PLRPGELLFRSDPLAYTVCKGSRGVVCDRCLLGKEKLM
AK010447: 1    MEALKVEKFTTANRGNGLRAVAPLRPGELLFRSDPLAYTVCKGSRGVVCDRCLLGKEKLM 60

ZNFN3A1 : 61   RCSQCRVAKYCSAKCQKKAWPDHKRECKCLKSCKPRYPPDSVRLLGRVVFKLMDGAPSES 120
               RCSQCR+AKYCSAKCQKKAWPDH+REC CLKSCKPRYPPDSVRLLGRV+ KLMD  PSES   B
AK010447: 61   RCSQCRIAKYCSAKCQKKAWPDHRRECSCLKSCKPRYPPDSVRLLGRVIVKLMDEKPSES 120

ZNFN3A1 : 121  EKLYSFYDLESNINKLTEDKKEGLRQLVMTFQHFMREEIQDASQLPPAFDLFEAFAKVIC 180
               EKLYSFYDLESNI+KLTEDKKEGLRQL MTFQHFMREEIQDASQLPP+FDLFEAFAKVIC
AK010447: 121  EKLYSFYDLESNISKLTEDKKEGLRQLAMTFQHFMREEIQDASQLPPSFDLFEAFAKVIC 180

ZNFN3A1 : 181  NSFTICNAEMQEVGVGLYPSISLLNHSCDPNCSIVFNGPHLLLRAVRDIEVGEELTICYL 240
               NSFTICNAEMQEVGVGLYPS+SLLNHSCDPNCSIVFNGPHLLLRAVR+IE GEELTICYL
AK010447: 181  NSFTICNAEMQEVGVGLYPSMSLLNHSCDPNCSIVFNGPHLLLRAVREIEAGEELTICYL 240

ZNFN3A1 : 241  DMLMTSEERRKQLRDQYCFECDCFRCQTQDKDADMLTGDEQVWKEVQESLKKIEELKAHW 300
               DMLMTSEERRKQLRDQYCFECDC RCQTQDKDADMLTGDEQ+WKEVQESLKKIEELKAHW
AK010447: 241  DMLMTSEERRKQLRDQYCFECDCIRCQTQDKDADMLTGDEQIWKEVQESLKKIEELKAHW 300

ZNFN3A1 : 301  KWEQVLAMCQAIISSNSERLPDINIYQLKVLDCAMDACINLGLLEEALFYGTRTMEPYRI 360
               KWEQVLA+CQAII+SNS RLPDINIYQLKVLDCAMDACINLG+LEEALFY  RTMEPYRI
AK010447: 301  KWEQVLALCQAIINSNSNRLPDINIYQLKVLDCAMDACINLGMLEEALFYAMRTMEPYRI 360

ZNFN3A1 : 361  FFPGSHPVRGVQVMKVGKLQLHQGMFPQAMKNLRLAFDIMRVTHGREHSLIEDLILLLEE 420
               FFPGSHPVRGVQVMKVGKLQLHQGMFPQAMKNLRLAFDIM+VTHGREHSLIEDLILLLEE
AK010447: 361  FFPGSHPVRGVQVMKVGKLQLHQGMFPQAMKNLRLAFDIMKVTHGREHSLIEDLILLLEE 420

ZNFN3A1 : 421  CDANIRAS 428
               CDANIRAS
AK010447: 421  CDANIRAS 428
```

FIG.3

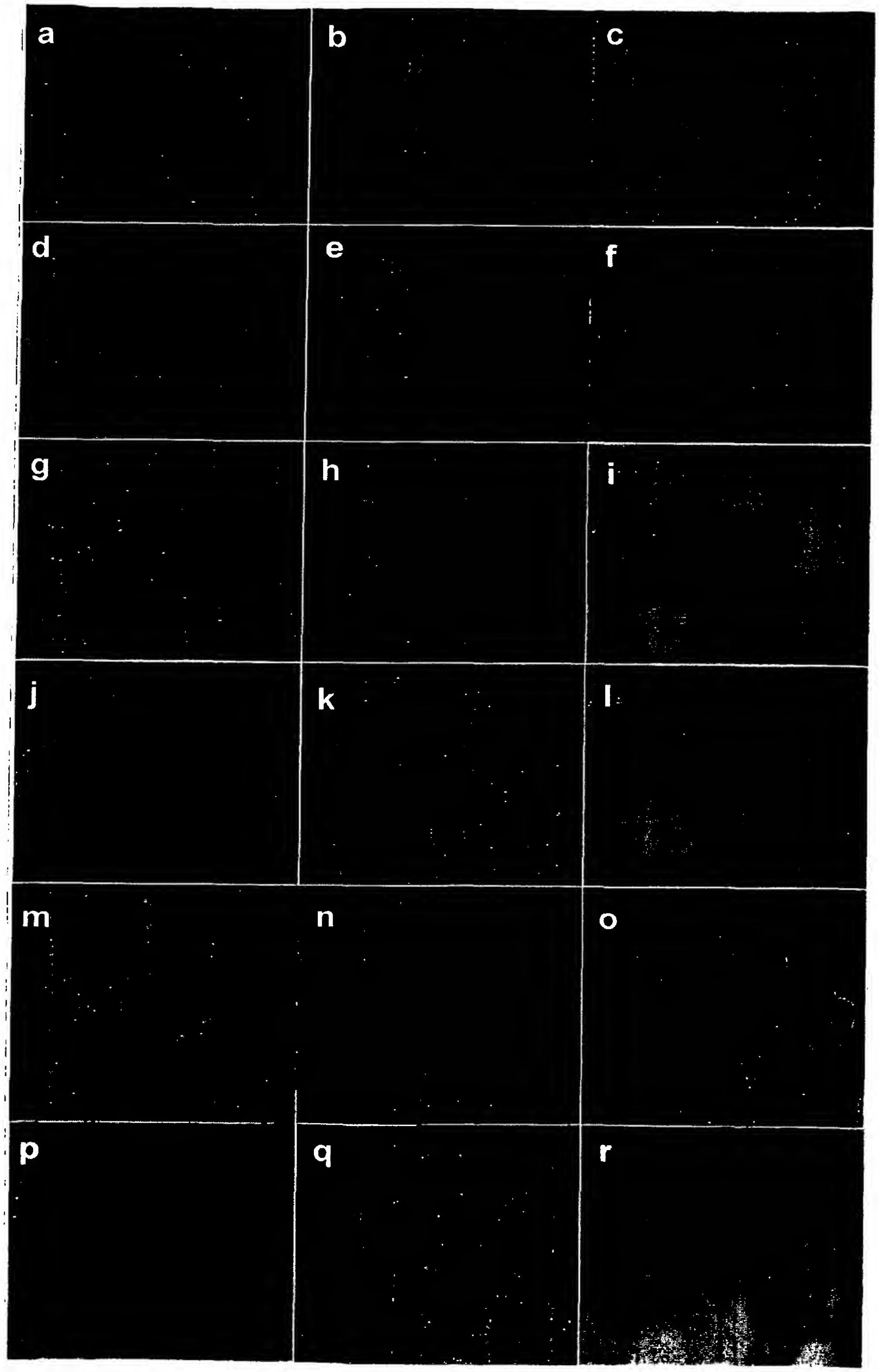
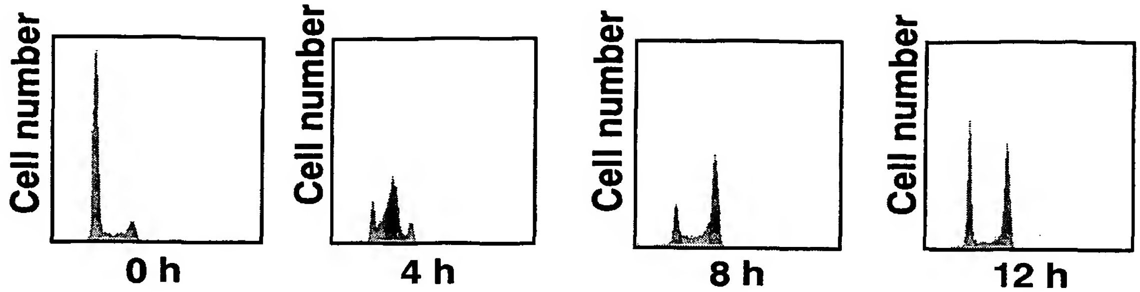


FIG.4

a


Cell number
Cell number
Cell number
Cell number
0 h
4 h
8 h
12 h

b

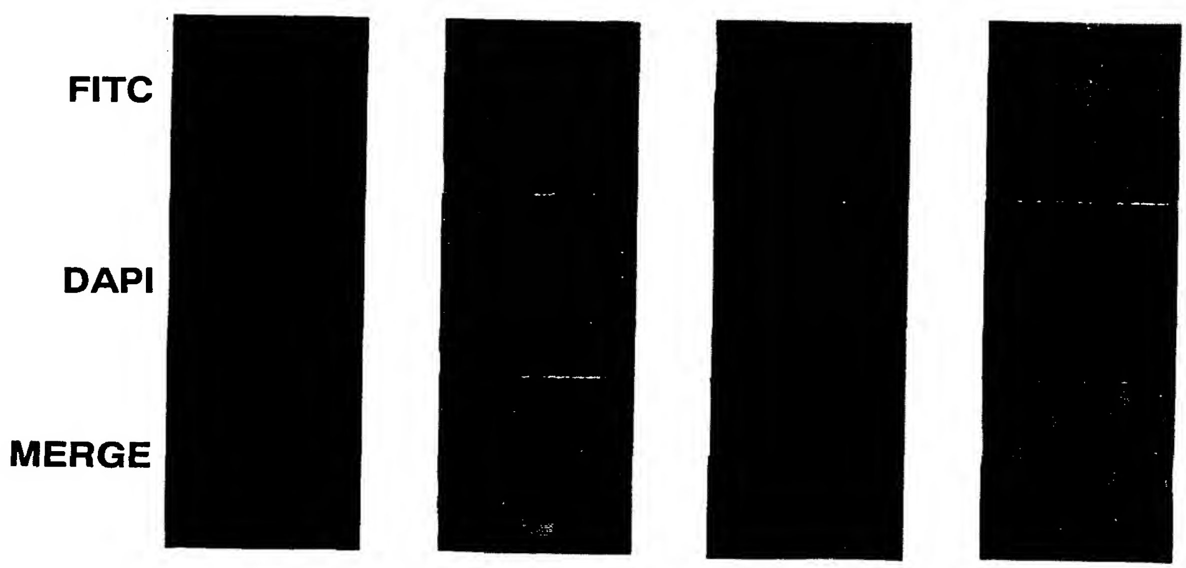

FITC
DAPI
MERGE

FIG.5

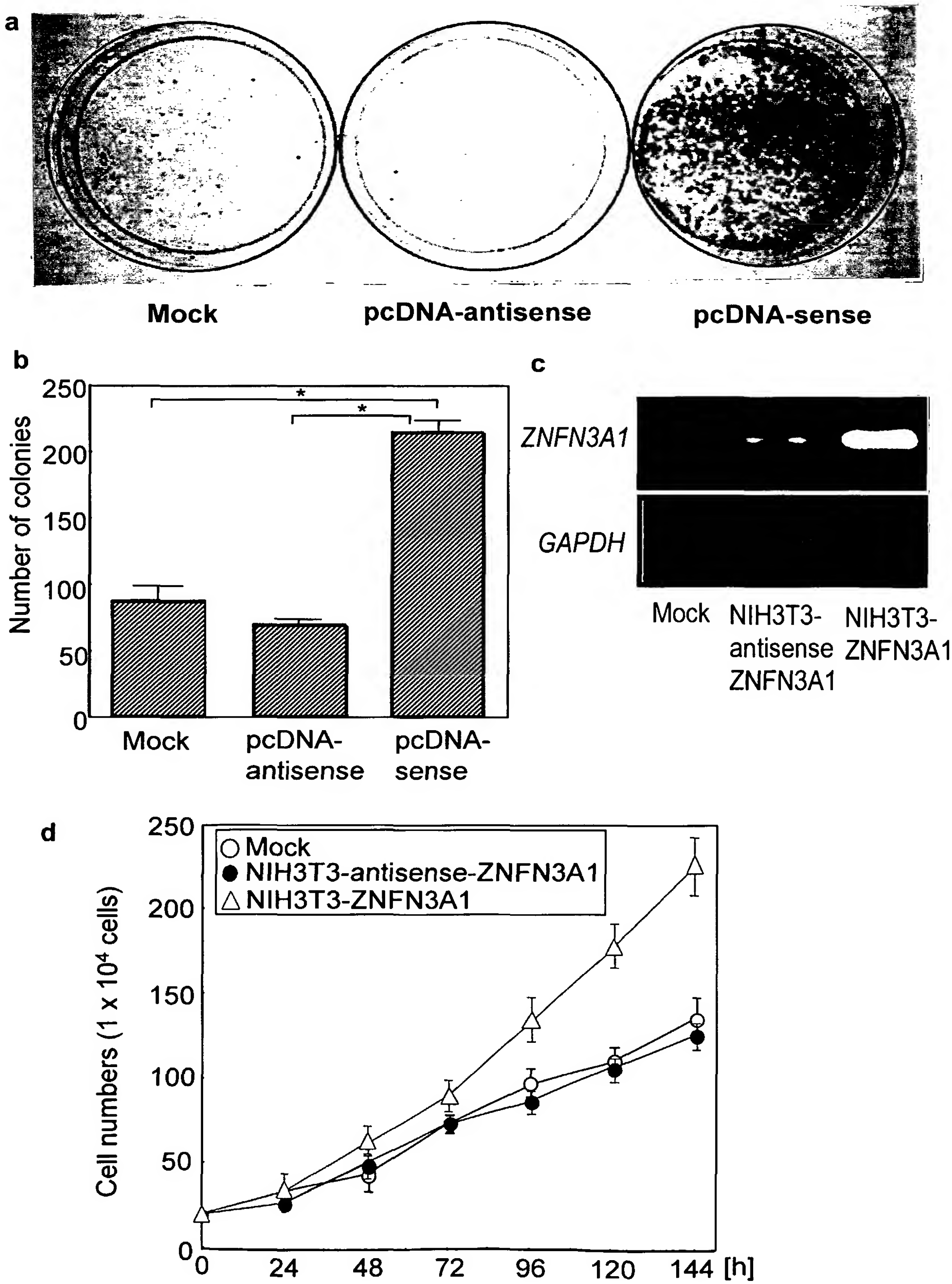

a
Mock
pcDNA-antisense
pcDNA-sense
b
Number of colonies
250
200
150
100
50
0
*
*
Mock
pcDNA-antisense
pcDNA-sense
c
ZNFN3A1
GAPDH
Mock
NIH3T3-antisense ZNFN3A1
NIH3T3-ZNFN3A1
d
Cell numbers (1 x 10^4 cells)
250
200
150
100
50
0
Mock
NIH3T3-antisense-ZNFN3A1
NIH3T3-ZNFN3A1
0
24
48
72
96
120
144 [h]

FIG.6

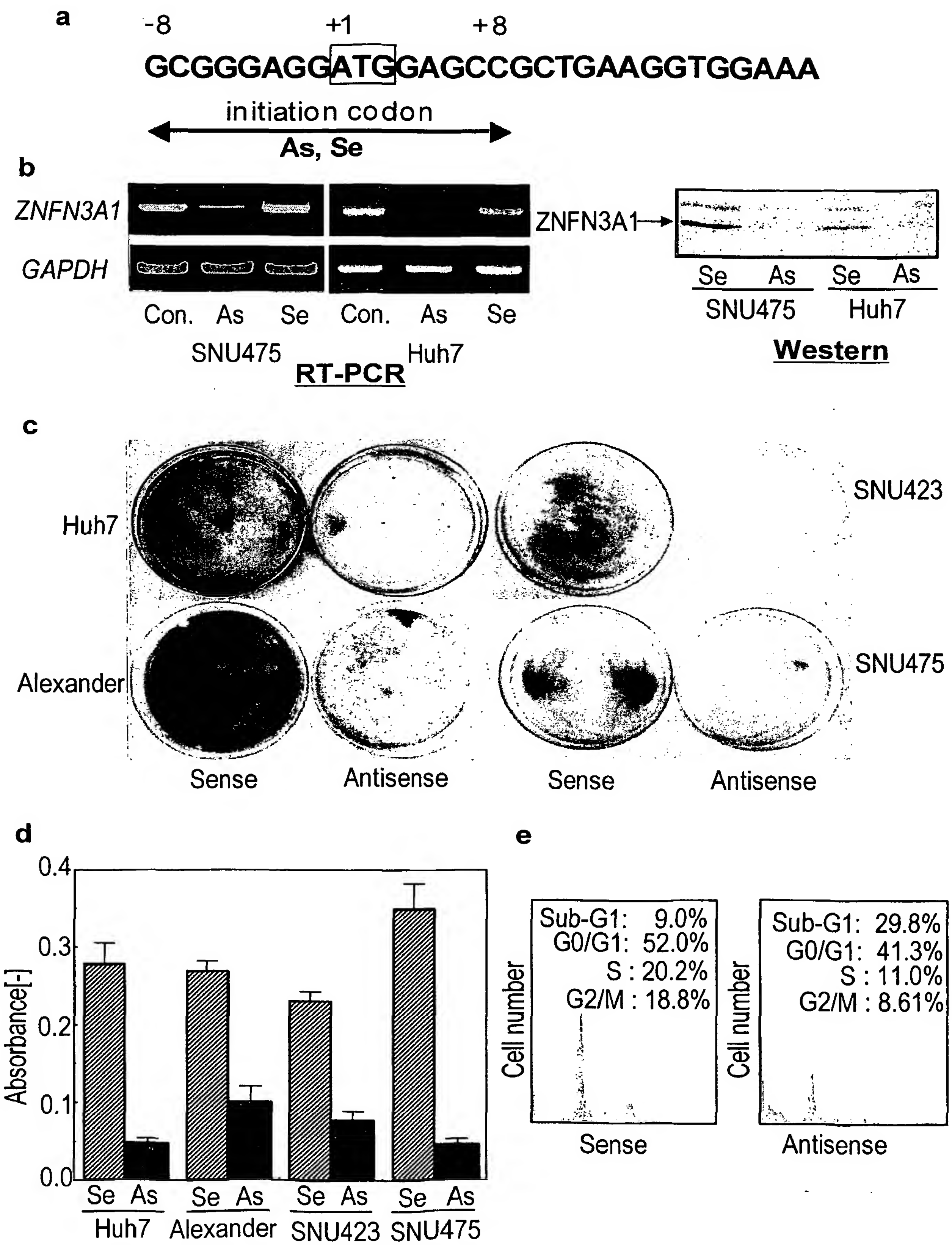

a
-8
+1
+8
GCGGGAGGATGGAGCCGCTGAAGGTGGAAA
initiation codon
As, Se
b
ZNFN3A1
GAPDH
Con.
As
Se
Con.
As
Se
SNU475
Huh7
RT-PCR
ZNFN3A1
Se
As
Se
As
SNU475
Huh7
Western
c
Huh7
Alexander
SNU423
SNU475
Sense
Antisense
Sense
Antisense
d
Absorbance[-]
0.4
0.3
0.2
0.1
0.0
Se
As
Huh7
Alexander
SNU423
SNU475
e
Sub-G1: 9.0%
G0/G1: 52.0%
S : 20.2%
G2/M : 18.8%
Cell number
Sense
Sub-G1: 29.8%
G0/G1: 41.3%
S : 11.0%
G2/M : 8.61%
Cell number
Antisense

FIG. 7

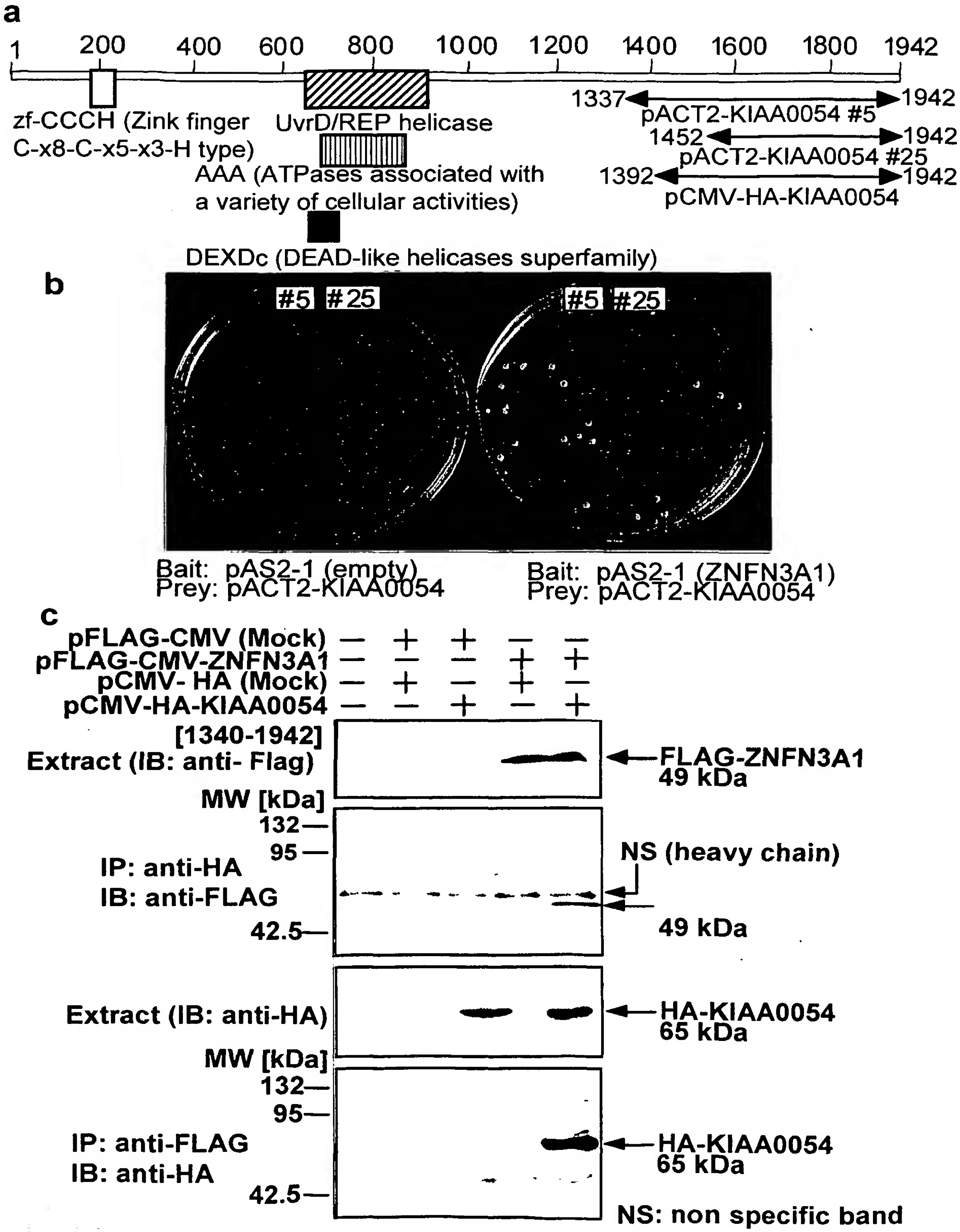

a
1
200
400
600
800
1000
1200
1400
1600
1800
1942
zf-CCCH (Zink finger C-x8-C-x5-x3-H type)
UvrD/REP helicase
AAA (ATPases associated with a variety of cellular activities)
DEXDc (DEAD-like helicases superfamily)
1337
1942
pACT2-KIAA0054 #5
1452
1942
pACT2-KIAA0054 #25
1392
1942
pCMV-HA-KIAA0054
b
#5
#25
#5
#25
Bait: pAS2-1 (empty)
Prey: pACT2-KIAA0054
Bait: pAS2-1 (ZNFN3A1)
Prey: pACT2-KIAA0054
c
pFLAG-CMV (Mock) − + + − −
pFLAG-CMV-ZNFN3A1 − − − + +
pCMV- HA (Mock) − + − + −
pCMV-HA-KIAA0054 − − + − +
[1340-1942]
Extract (IB: anti- Flag)
FLAG-ZNFN3A1
49 kDa
MW [kDa]
132
95
42.5
IP: anti-HA
IB: anti-FLAG
NS (heavy chain)
49 kDa
Extract (IB: anti-HA)
HA-KIAA0054
65 kDa
MW [kDa]
132
95
42.5
IP: anti-FLAG
IB: anti-HA
HA-KIAA0054
65 kDa
NS: non specific band

FIG. 8

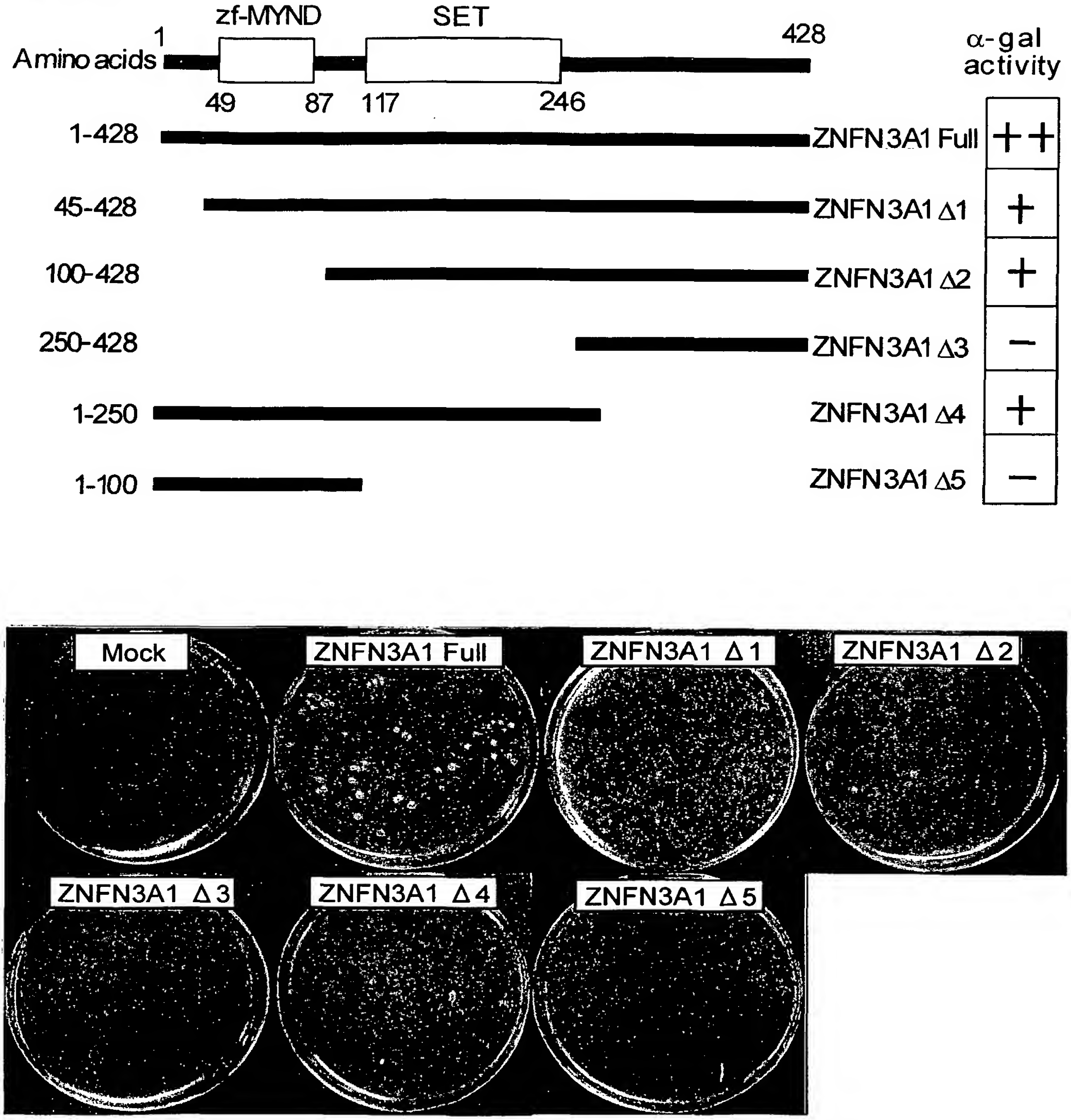

zf-MYND
SET
Amino acids
1
428
49
87
117
246
α-gal activity
1-428
ZNFN3A1 Full
++
45-428
ZNFN3A1 Δ1
+
100-428
ZNFN3A1 Δ2
+
250-428
ZNFN3A1 Δ3
−
1-250
ZNFN3A1 Δ4
+
1-100
ZNFN3A1 Δ5
−
Mock
ZNFN3A1 Full
ZNFN3A1 Δ1
ZNFN3A1 Δ2
ZNFN3A1 Δ3
ZNFN3A1 Δ4
ZNFN3A1 Δ5

FIG. 9

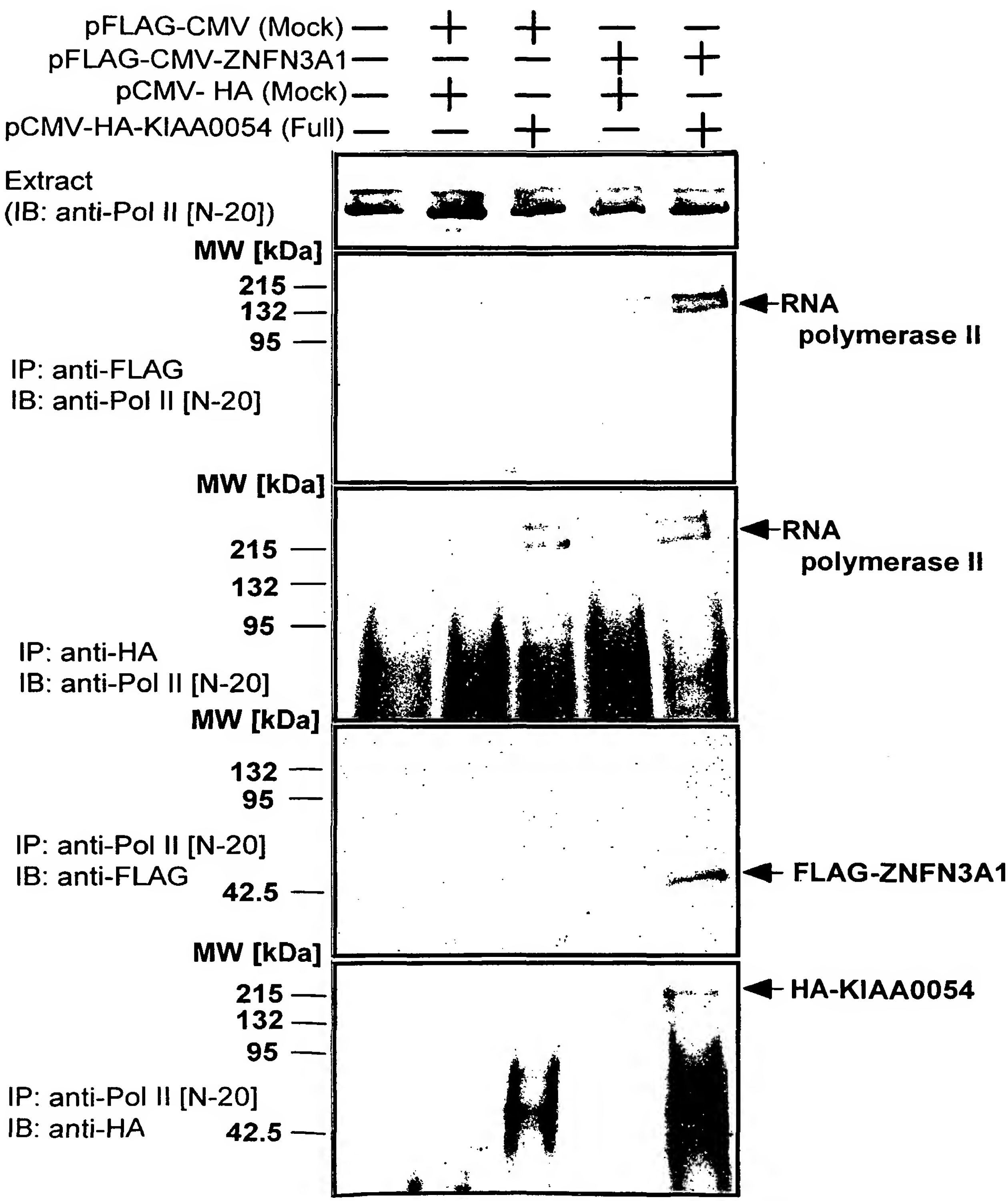

pFLAG-CMV (Mock) — + + — —
pFLAG-CMV-ZNFN3A1 — — — + +
pCMV- HA (Mock) — + — + —
pCMV-HA-KIAA0054 (Full) — — + — +
Extract
(IB: anti-Pol II [N-20])
MW [kDa]
215
132
95
IP: anti-FLAG
IB: anti-Pol II [N-20]
RNA polymerase II
MW [kDa]
215
132
95
IP: anti-HA
IB: anti-Pol II [N-20]
RNA polymerase II
MW [kDa]
132
95
IP: anti-Pol II [N-20]
IB: anti-FLAG
42.5
FLAG-ZNFN3A1
MW [kDa]
215
132
95
IP: anti-Pol II [N-20]
IB: anti-HA
42.5
HA-KIAA0054

FIG. 10

a

	Experimental probability	Theoretical probability	Experimental / Theoretical
C C C T C	(37.0%)	0.015 (1.5%)	24.6
C C C T C C	(34.8%)	0.0034 (0.34%)	102.3
C C C C T C C	(20.7%)	0.00079 (0.079%)	262
C C C T C C T	(20.0%)	0.00079 (0.079%)	253
(C) C C C T C C (T)	(10.6%)	0.00018 (0.018%)	589
(A) G G A G G G (G)	(10.6%)	0.00018 (0.018%)	589

b

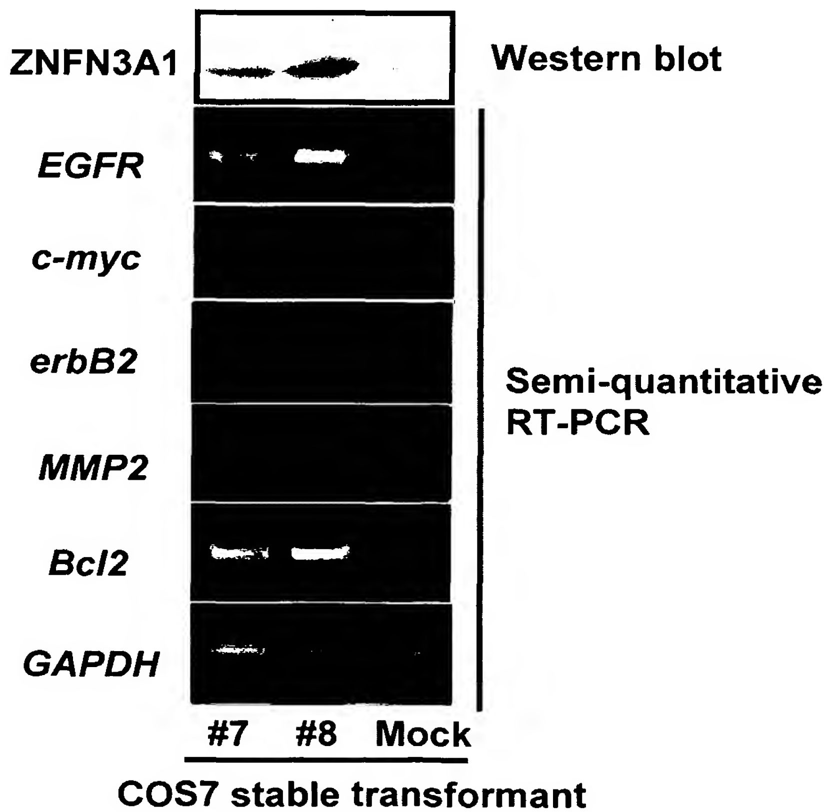


FIG. 11

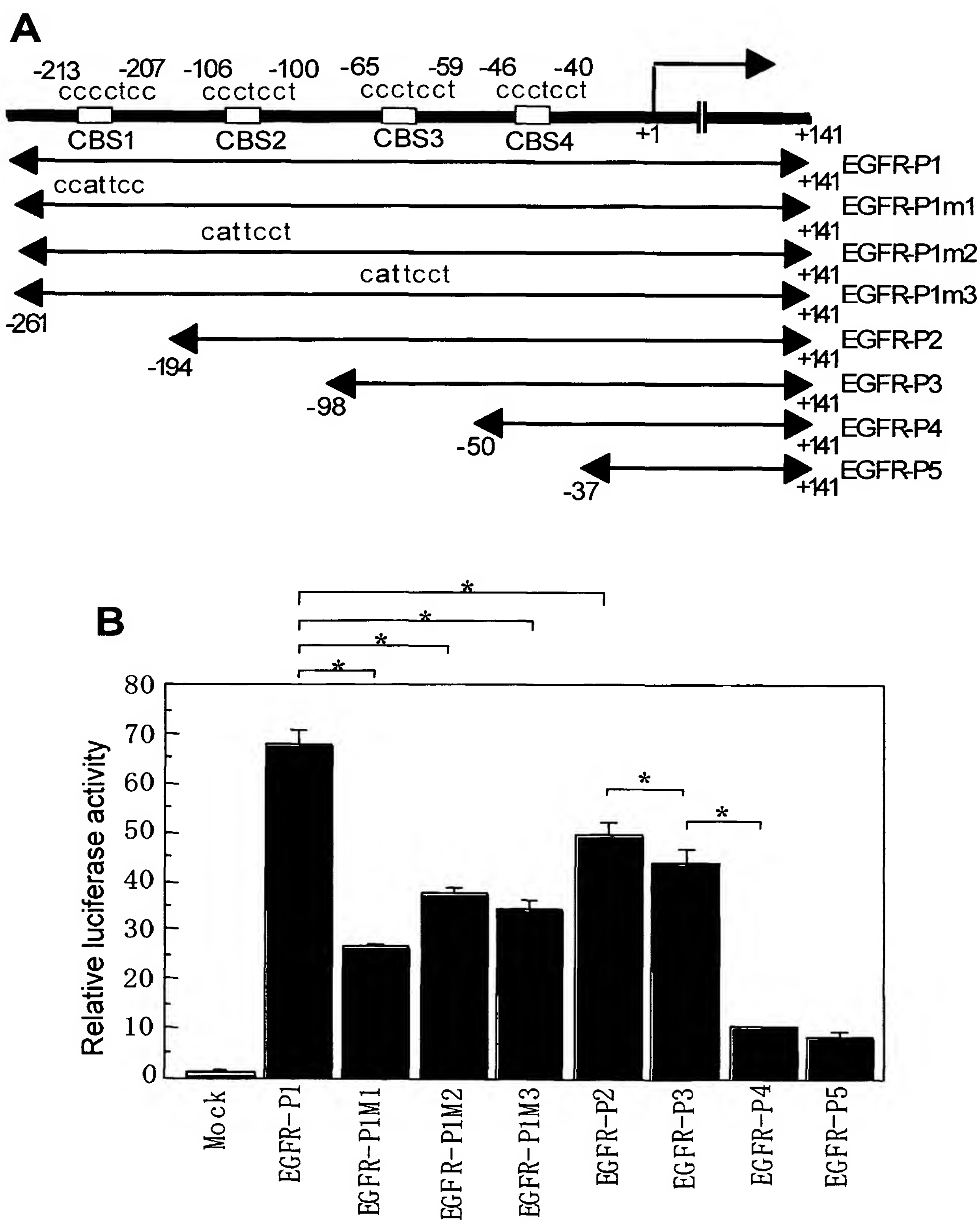
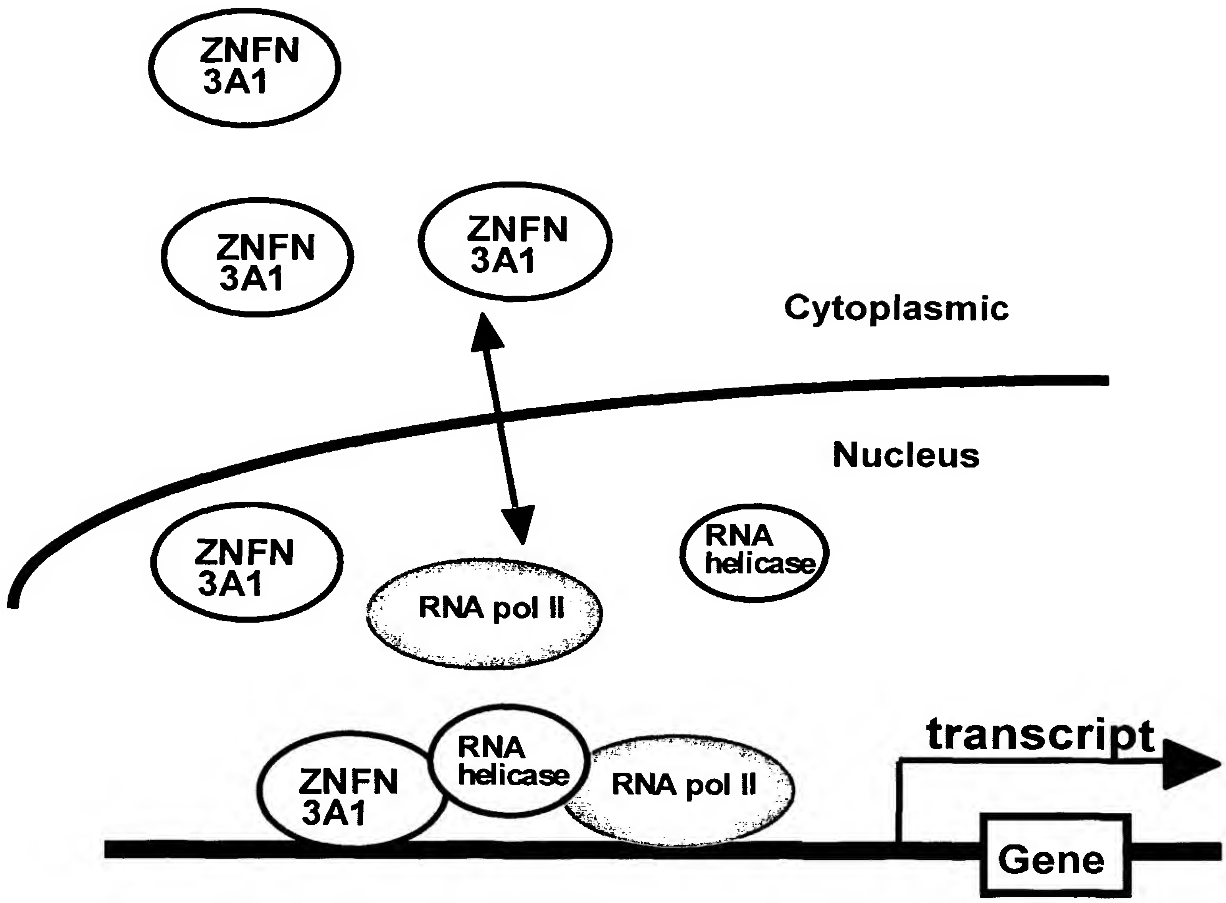

A
-213 -207 -106 -100 -65 -59 -46 -40
cccctcc
ccctcct
ccctcct
ccctcct
CBS1
CBS2
CBS3
CBS4
+1
+141
EGFR-P1
ccattcc
EGFR-P1m1
cattcct
EGFR-P1m2
cattcct
EGFR-P1m3
-261
EGFR-P2
-194
EGFR-P3
-98
EGFR-P4
-50
EGFR-P5
-37
B
Relative luciferase activity
0
10
20
30
40
50
60
70
80
*
Mock
EGFR-P1
EGFR-P1M1
EGFR-P1M2
EGFR-P1M3
EGFR-P2
EGFR-P3
EGFR-P4
EGFR-P5

FIG. 12


ZNFN 3A1
ZNFN 3A1
ZNFN 3A1
Cytoplasmic
Nucleus
ZNFN 3A1
RNA helicase
RNA pol II
transcript
ZNFN 3A1
RNA helicase
RNA pol II
Gene

SEQUENCE LISTING

<110> JAPAN AS REPRESENTED BY THE PRESIDENT OF THE UNIVERSITY OF TOKYO
ONCOTHERAPY SCIENCE, INC.

<120> Gene and Protein Relating To Hepatocellular Carcinoma

<130> ONC-A0206P

<150> US 60/324.261
<151> 2001-09-25

<150> US 60/391,666
<151> 2002-06-26

<150> CA
<151> 2002-08-23

<160> 17

<170> PatentIn version 3.1

<210> 1
<211> 1622
<212> DNA
<213> Homo sapiens

<220>

<221> CDS

<222> (96)..(1382)

<223>

<400> 1

gtgcgcgcag ggcgcaggcg cgcgggtccc ggcagcccgt gagacgcccg ctgctggacg 60

cgggtagccg tctgaggtgc cggagctgcg ggagg atg gag ccg ctg aag gtg 113
Met Glu Pro Leu Lys Val
1 5

gaa aag ttc gca acc gcc aac agg gga aac ggg ctg cgc gcc gtg acc 161
Glu Lys Phe Ala Thr Ala Asn Arg Gly Asn Gly Leu Arg Ala Val Thr
10 15 20

ccg ctg cgc ccc gga gag cta ctc ttc cgc tcg gat ccc ttg gcg tac 209
Pro Leu Arg Pro Gly Glu Leu Leu Phe Arg Ser Asp Pro Leu Ala Tyr
25 30 35

acg gtg tgc aag ggg agt cgt ggc gtc gtc tgc gac cgc tgc ctt ctc 257
Thr Val Cys Lys Gly Ser Arg Gly Val Val Cys Asp Arg Cys Leu Leu
40 45 50

ggg aag gaa aag ctg atg cga tgc tct cag tgc cgc gtc gcc aaa tac 305

Gly Lys Glu Lys Leu Met Arg Cys Ser Gln Cys Arg Val Ala Lys Tyr
55 60 65 70

tgt agt gct aag tgt cag aaa aaa gct tgg cca gac cac aag cgg gaa 353
Cys Ser Ala Lys Cys Gln Lys Lys Ala Trp Pro Asp His Lys Arg Glu
75 80 85

tgc aaa tgc ctt aaa agc tgc aaa ccc aga tat cct cca gac tcc gtt 401
Cys Lys Cys Leu Lys Ser Cys Lys Pro Arg Tyr Pro Pro Asp Ser Val
90 95 100

cga ctt ctt ggc aga gtt gtc ttc aaa ctt atg gat gga gca cct tca 449
Arg Leu Leu Gly Arg Val Val Phe Lys Leu Met Asp Gly Ala Pro Ser
105 110 115

gaa tca gag aag ctt tac tca ttt tat gat ctg gag tca aat att aac 497
Glu Ser Glu Lys Leu Tyr Ser Phe Tyr Asp Leu Glu Ser Asn Ile Asn
120 125 130

aaa ctg act gaa gat aag aaa gag ggc ctc agg caa ctc gta atg aca 545
Lys Leu Thr Glu Asp Lys Lys Glu Gly Leu Arg Gln Leu Val Met Thr
135 140 145 150

ttt caa cat ttc atg aga gaa gaa ata cag gat gcc tct cag ctg cca 593
Phe Gln His Phe Met Arg Glu Glu Ile Gln Asp Ala Ser Gln Leu Pro
155 160 165

cct gcc ttt gac ctt ttt gaa gcc ttt gca aaa gtg atc tgc aac tct 641
Pro Ala Phe Asp Leu Phe Glu Ala Phe Ala Lys Val Ile Cys Asn Ser
170 175 180

ttc acc atc tgt aat gcg gag atg cag gaa gtt ggt gtt ggc cta tat 689
Phe Thr Ile Cys Asn Ala Glu Met Gln Glu Val Gly Val Gly Leu Tyr
185 190 195

ccc agt atc tct ttg ctc aat cac agc tgt gac ccc aac tgt tcg att 737
Pro Ser Ile Ser Leu Leu Asn His Ser Cys Asp Pro Asn Cys Ser Ile
200 205 210

gtg ttc aat ggg ccc cac ctc tta ctg cga gca gtc cga gac atc gag 785
Val Phe Asn Gly Pro His Leu Leu Leu Arg Ala Val Arg Asp Ile Glu
215 220 225 230

gtg gga gag gag ctc acc atc tgc tac ctg gat atg ctg atg acc agt 833
Val Gly Glu Glu Leu Thr Ile Cys Tyr Leu Asp Met Leu Met Thr Ser
235 240 245

gag gag cgc cgg aag cag ctg agg gac cag tac tgc ttt gaa tgt gac 881
Glu Glu Arg Arg Lys Gln Leu Arg Asp Gln Tyr Cys Phe Glu Cys Asp
250 255 260

tgt ttc cgt tgc caa acc cag gac aag gat gct gat atg cta act ggt 929

Cys Phe Arg Cys Gln Thr Gln Asp Lys Asp Ala Asp Met Leu Thr Gly
265 270 275

gat gag caa gta tgg aag gaa gtt caa gaa tcc ctg aaa aaa att gaa 977
Asp Glu Gln Val Trp Lys Glu Val Gln Glu Ser Leu Lys Lys Ile Glu
280 285 290

gaa ctg aag gca cac tgg aag tgg gag cag gtt ctg gcc atg tgc cag 1025
Glu Leu Lys Ala His Trp Lys Trp Glu Gln Val Leu Ala Met Cys Gln
295 300 305 310

gcg atc ata agc agc aat tct gaa cgg ctt ccc gat atc aac atc tac 1073
Ala Ile Ile Ser Ser Asn Ser Glu Arg Leu Pro Asp Ile Asn Ile Tyr
315 320 325

cag ctg aag gtg ctc gac tgc gcc atg gat gcc tgc atc aac ctc ggc 1121
Gln Leu Lys Val Leu Asp Cys Ala Met Asp Ala Cys Ile Asn Leu Gly
330 335 340

ctg ttg gag gaa gcc ttg ttc tat ggt act cgg acc atg gag cca tac 1169
Leu Leu Glu Glu Ala Leu Phe Tyr Gly Thr Arg Thr Met Glu Pro Tyr
345 350 355

agg att ttt ttc cca gga agc cat ccc gtc aga ggg gtt caa gtg atg 1217
Arg Ile Phe Phe Pro Gly Ser His Pro Val Arg Gly Val Gln Val Met
360 365 370

aaa gtt ggc aaa ctg cag cta cat caa ggc atg ttt ccc caa gca atg 1265
Lys Val Gly Lys Leu Gln Leu His Gln Gly Met Phe Pro Gln Ala Met
375 380 385 390

aag aat ctg aga ctg gct ttt gat att atg aga gtg aca cat ggc aga 1313
Lys Asn Leu Arg Leu Ala Phe Asp Ile Met Arg Val Thr His Gly Arg
395 400 405

gaa cac agc ctg att gaa gat ttg att cta ctt tta gaa gaa tgc gac 1361
Glu His Ser Leu Ile Glu Asp Leu Ile Leu Leu Leu Glu Glu Cys Asp
410 415 420

gcc aac atc aga gca tcc taa gggaacgcag tcagagggaa atacggcgtg 1412
Ala Asn Ile Arg Ala Ser
425

tgtctttgtt gaatgcctta ttgaggtcac acactctatg ctttgttagc tgtgtgaacc 1472

tctcttattg gaaattctgt tccgtgtttg tgtaggtaaa taaaggcaga catggtttgc 1532

aaaccacaag aatcattagt tgtagagaag cacgattata ataaattcaa aacatttggt 1592

tgaggatgcc aaaaaaaaaa aaaaaaaaaa 1622

<210> 2

<211> 428

<212> PRT

<213> Homo sapiens

<400> 2

Met Glu Pro Leu Lys Val Glu Lys Phe Ala Thr Ala Asn Arg Gly Asn
1 5 10 15

Gly Leu Arg Ala Val Thr Pro Leu Arg Pro Gly Glu Leu Leu Phe Arg
20 25 30

Ser Asp Pro Leu Ala Tyr Thr Val Cys Lys Gly Ser Arg Gly Val Val
35 40 45

Cys Asp Arg Cys Leu Leu Gly Lys Glu Lys Leu Met Arg Cys Ser Gln
50 55 60

Cys Arg Val Ala Lys Tyr Cys Ser Ala Lys Cys Gln Lys Lys Ala Trp
65 70 75 80

Pro Asp His Lys Arg Glu Cys Lys Cys Leu Lys Ser Cys Lys Pro Arg
85 90 95

Tyr Pro Pro Asp Ser Val Arg Leu Leu Gly Arg Val Val Phe Lys Leu
100 105 110

Met Asp Gly Ala Pro Ser Glu Ser Glu Lys Leu Tyr Ser Phe Tyr Asp
115 120 125

Leu Glu Ser Asn Ile Asn Lys Leu Thr Glu Asp Lys Lys Glu Gly Leu
130 135 140

Arg Gln Leu Val Met Thr Phe Gln His Phe Met Arg Glu Glu Ile Gln
145 150 155 160

Asp Ala Ser Gln Leu Pro Pro Ala Phe Asp Leu Phe Glu Ala Phe Ala
165 170 175

Lys Val Ile Cys Asn Ser Phe Thr Ile Cys Asn Ala Glu Met Gln Glu
180 185 190

Val Gly Val Gly Leu Tyr Pro Ser Ile Ser Leu Leu Asn His Ser Cys
195 200 205

Asp Pro Asn Cys Ser Ile Val Phe Asn Gly Pro His Leu Leu Leu Arg
210 215 220

Ala Val Arg Asp Ile Glu Val Gly Glu Glu Leu Thr Ile Cys Tyr Leu
225 230 235 240

Asp Met Leu Met Thr Ser Glu Glu Arg Arg Lys Gln Leu Arg Asp Gln

245 250 255

Tyr Cys Phe Glu Cys Asp Cys Phe Arg Cys Gln Thr Gln Asp Lys Asp

260 265 270

Ala Asp Met Leu Thr Gly Asp Glu Gln Val Trp Lys Glu Val Gln Glu

275 280 285

Ser Leu Lys Lys Ile Glu Glu Leu Lys Ala His Trp Lys Trp Glu Gln

290 295 300

Val Leu Ala Met Cys Gln Ala Ile Ile Ser Ser Asn Ser Glu Arg Leu

305 310 315 320

Pro Asp Ile Asn Ile Tyr Gln Leu Lys Val Leu Asp Cys Ala Met Asp

325 330 335

Ala Cys Ile Asn Leu Gly Leu Leu Glu Glu Ala Leu Phe Tyr Gly Thr

340 345 350

Arg Thr Met Glu Pro Tyr Arg Ile Phe Phe Pro Gly Ser His Pro Val

355 360 365

Arg Gly Val Gln Val Met Lys Val Gly Lys Leu Gln Leu His Gln Gly

370 375 380

Met Phe Pro Gln Ala Met Lys Asn Leu Arg Leu Ala Phe Asp Ile Met
385 390 395 400

Arg Val Thr His Gly Arg Glu His Ser Leu Ile Glu Asp Leu Ile Leu
405 410 415

Leu Leu Glu Glu Cys Asp Ala Asn Ile Arg Ala Ser
420 425

<210> 3
<211> 16
<212> DNA
<213> Artificial

<220>
<223> an artificially synthesized primer sequence

<400> 3
gcgggaggat ggagcc 16

<210> 4
<211> 22
<212> DNA
<213> Artificial

<220>

<223> an artificially synthesized primer sequence

<400> 4

acaacagcct caagatcatc ag 22

<210> 5

<211> 20

<212> DNA

<213> Artificial

<220>

<223> an artificially synthesized primer sequence

<400> 5

ggtccaccac tgacacgttg 20

<210> 6

<211> 23

<212> DNA

<213> Artificial

<220>

<223> an artificially synthesized primer sequence

<400> 6
ttcccgatat caacatctac cag 23

<210> 7
<211> 23
<212> DNA
<213> Artificial

<220>
<223> an artificially synthesized primer sequence

<400> 7
agtgtgtgac ctcaataagg cat 23

<210> 8
<211> 24
<212> DNA
<213> Artificial

<220>
<223> an artificially synthesized primer sequence

<400> 8

ctgccaagaa gtcggagtct ggag 24

<210> 9

<211> 32

<212> DNA

<213> Artificial

<220>

<223> an artificially synthesized primer sequence

<400> 9

cggaattcat ggagccgctg aaggtggaaa ag 32

<210> 10

<211> 31

<212> DNA

<213> Artificial

<220>

<223> an artificially synthesized primer sequence

<400> 10

ccgctcgagg gatgctctga tgttggcgtc g 31

<210> 11

<211> 64

<212> DNA

<213> Artificial

<220>

<223> artificially synthesized oligonucleotides

<220>

<221> misc_feature

<222> (21)..(40)

<223> "n"=A, G, C or T

<400> 11

gggagaattc cgacacgcgt nnnnnnnnnn nnnnnnnnnn ctcgagcgtc tacatggatc 60

ctca 64

<210> 12

<211> 31

<212> DNA

<213> Artificial

<220>

<223> an artificially synthesized primer sequence

<400> 12

ggggtaccca gtgctgggaa cgcccctctc g 31

<210> 13

<211> 31

<212> DNA

<213> Artificial

<220>

<223> an artificially synthesized primer sequence

<400> 13

ggggtaccca ctcccgccgg agactaggtc c 31

<210> 14

<211> 31

<212> DNA

<213> Artificial

<220>

<223> an artificially synthesized primer sequence

<400> 14

ggggtaccct cgcattctcc tcctcctctg c 31

<210> 15

<211> 31

<212> DNA

<213> Artificial

<220>

<223> an artificially synthesized primer sequence

<400> 15

ggggtacctg gtccctcctc ctcccgccct g 31

<210> 16

<211> 31

<212> DNA

<213> Artificial

<220>

<223> an artificially synthesized primer sequence

<400> 16

ggggtacctc ccgccctgcc tcccgcgcct c 31

<210> 17

<211> 31

<212> DNA

<213> Artificial

<220>

<223> an artificially synthesized primer sequence

<400> 17

gaagatctag gtggcctgtc gtccggtctg g 31